

sequence search results

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:20:14 ; Search time 19.5 Seconds
(without alignments)
6241.288 Million cell updates/sec

Title: US-09-714-936-218

Perfect score: 2383
Sequence: 1 ccggaattccgggtcgagc.....tttctctctctttttttg 1294

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US0714936/runat_10092003_151458_5680/app_query_fasta_1.1479
-DB=SwissProt_41 -OFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0714936.cgn_1_1 -runat_10092003_151458_5680 -NCPU=6 -ICPU=3
-NO_MMAP -IARGUMENT -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	941	39.5	305	1 SI7C_RAT	Q64686 rattus norv
2	917	38.5	305	1 SI7C_MOUSE	Q9WUV2 mus musculu
3	415	17.4	360	1 SI7D_MOUSE	Q9R2B6 mus musculu
4	382	16.0	302	1 SI7D_HUMAN	Q9H4F1 h alpha-n-a
5	328.5	13.8	336	1 SI7E_HUMAN	Q9BVH7 homo sapien
6	319.5	13.4	336	1 SI7E_MOUSE	Q9QYJ1 mus musculu
7	163	6.8	413	1 SI7A_CHICK	Q921B2 gallus gall
8	146.5	6.1	403	1 SI7A_MOUSE	Q64685 mus musculu
9	142	6.0	406	1 SI7A_HUMAN	P15097 homo sapien
10	138.5	5.8	374	1 SI7A_RAT	Q02734 r rattus norv
11	138.5	5.8	403	1 SI7A_MOUSE	P97325 mus musculu
12	137.5	5.8	374	1 SI7A_HUMAN	Q11203 homo sapien
13	137.5	5.7	342	1 SI7A_CHICK	Q11200 g cmp-n-ace
14	136	5.7	342	1 SI7A_MOUSE	Q11204 m cmp-n-ace
15	135.5	5.7	350	1 SI7B_MOUSE	Q11205 r cmp-n-ace
16	135.5	5.7	350	1 SI7B_RAT	Q11201 h cmp-n-ace
17	134	5.6	340	1 SI7B_HUMAN	Q16842 h cmp-n-ace
18	131.5	5.5	350	1 SI7B_MOUSE	

ALIGNMENTS

RESULT 1

SI7C_RAT	SI7C_RAT	STANDARD;	PRT;	305 AA.
ID	SI7C_RAT	STANDARD;	PRT;	305 AA.
AC	Q64686;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase III			
DE	(EC 2.4.99.-) (St6GalNAc III) (Sialyltransferase 7C) (STT).			
GN	STAT7C.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-Sprague-Dawley; TISSUE=Brain;			
RX	MEDLINE=96205924; PubMed=8631773;			
RA	Stoberg E.R., Kitagawa H., Glushka J., van Halbeek H., Paulson J.C.;			
RT	"Molecular cloning of a developmentally regulated			
RT	N-acetylgalactosamine alpha2,6-sialyltransferase specific for			
RL	sialylated glycoconjugates";			
RL	J. Biol. Chem. 271:7450-7459(1996).			
CC	-I- CATALYTIC ACTIVITY: CMP-N-acetylneuraminic acid + alpha-N-			
CC	acetylneuraminyl-2,3-beta-D-galactosyl-1,3-(N-acetyl-D-			
CC	galactosaminyl)-glycoprotein = CMP + alpha-N-acetylneuraminyl-			
CC	2,3-beta-D-galactosyl-1,3-(N-acetylneuraminyl)-2,6-alpha-N-acetyl-			
CC	neuraminyl)-glycoprotein.			
CC	-I- PATHWAY: Glycosylation.			
CC	-I- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).			
CC	-I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, NEWBORN BRAIN AND			
CC	KIDNEY AND TO A LESSER EXTENT IN LUNG. NOT FOUND IN LIVER AND			
CC	SKELETAL MUSCLE.			
CC	-I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

```

CC EMBL; L29554; AAC2086.1; -.
DR InterPro; IPR001675; Glyco.trans_29.
DR Pfam; PF00777; Glyco.transf.29; 1.
KW transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 9 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 30 305 LUMENAL, CATALYTIC (POTENTIAL).
FT DISULFID 80 229 BY SIMILARITY.
FT CARBOHYD 239 239 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 305 AA; 35149 MW; 9077F6547D359AE9 CRC64;

Alignment Scores:
Pred. No.: 3,36e-80 Length: 305
Score: 941.00 Matches: 173
Percent Similarity: 91.35% Conservative: 17
Best Local Similarity: 83.17% Mismatches: 18
Query Match: 39.49% Indels: 0
DB: 1 Gaps: 0

US-09-714-936-218 (1-1294) x SI7C_RAT (1-305)
QY 166 ATGCGCTGATCTCGAGAGAACTGCTGATGCTGAGCTCATAGCAGCGTTCCT 225
Db 1 MetAlaCysIleuLeuYsArgProAlaLeuAlaValSerPheIleAlaLeuCysIle 20
QY 226 TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
Db 21 LeuLeuLeuAlaMetArgLeuAlaAsnAspValThrPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACTGGTACAAAGTGATGATCTTCCTACATACAGCGCGCCCTTCGAACT 345
Db 41 GlyGlnProYsThrIleYsPheProLeuSerTyThrLeuArgGlnProLeuGlnThr 60
QY 346 CACATGATGATACATAATGTGAAGACACAGAGCGCTTGCACCTGCTGACCTTGT 405
Db 61 HisTyGlyTyrIleAsnValArgThrGlnGluProLeuGlnLeuAsnCysAsnHisCys 80
QY 406 GCCATAGTCTCAACATCAGGTCAGATGTTGGCCAGAGGTGGGAATGAGATAGATCGA 465
Db 81 AlaValValSerAsnSerGlyGlnMetValGlyGlnValGlyGlnValGlyGlnVal 100
QY 466 TCCCTCCGATTTGGAGATGACAAATGCCGCCACCAAGGTTATGAGAGATGTCGGC 525
Db 101 AlaSerCysIleIlePargMetAsnAsnAlaProThrIleGlyPheGluGluValGly 120
QY 526 CGCATGACCATGATTCGAGTTGTGTCCTCATCAGCGCTTCCTCTTTTCTAAAAACCT 585
Db 121 TyrMetThrMetValArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATATTTTTCAGGAACGCAATACACTACTATTGTGTTATTTGGGGACCTTCCCGAAT 645
Db 141 AspTyrPhePheYsGluAlaSerThrThrIleTyrValIleTyrPheArgAsn 160
QY 646 ATGAGAAAGATGCAATGGCATGCTTTACACATGTTGAAAAGACAGTTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysThrValAspAlaTyr 180
QY 706 CGGAATGCCCAATATACGTGACCAACAGAGAGCGCATGAGTTACTGTGTGGATTTTT 765
Db 181 ProAspAlaGlnIleTyrValThrThrGluGlnArgMetThrTyrCysAspGlyValPhe 200
QY 766 AGAAGAGAACTGGGAGACAGG 789
Db 201 LysAspGluThrGlyLysAspArg 208

RESULT 2
SI7C_MOUSE STANDARD; PRT; 305 AA.
ID SI7C_MOUSE
AC Q9WU2; Q9JHP5;

```

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase III
DE (EC 2.4.99.-) (St6GalNac III) (Sialyltransferase 7C) (STY).
GN SIAT7C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
STAGE.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=9923522; PubMed=10207017;
RA Lee Y.-C., Kaufman M., Kitazume-Kawaguchi S., Kono M., Takashima S.,
RA Kurosawa N., Liu H., Pircher H., Tsuji S.;
RT "Molecular cloning and functional expression of two members of mouse
RT NeuAc2alpha2,3Galbeta1,3GalNac GalNAc2alpha2,6-sialyltransferase family,
RT St6GalNac III and IV."
RL J. Biol. Chem. 274:11958-11967(1999).
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF GANGLIOSIDE GD1A FROM
CC GM1B. TRANSFERS CMP-NEUAC WITH AN ALPHA-2,6-LINKAGE TO GALNAC
CC RESIDUE ON NEUAC-ALPHA-2,3-GAL-BETA-1,3-GALNAC OF GLYCOPROTEINS
CC AND GLYCOLIPIDS. ST6GALNACIII PREFERS GLYCOLIPIDS TO
CC GLYCOPROTEINS.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminyl + alpha-N-
CC acetylneuraminyl-2,3-beta-D-galactosyl-1,3-(N-acetyl-D-
CC galactosaminyl)-glycoprotein = CMP + alpha-N-acetylneuraminyl-
CC 2,3-beta-D-galactosyl-(2,6-alpha-N-acetylneuraminyl)-(N-acetyl-D-
CC galactosaminyl)-glycoprotein.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- TISSUE SPECIFICITY: HIGH EXPRESSION IN BRAIN, LUNG AND HEART AND
CC TO A LESSER EXTENT IN KIDNEY, MAMMARY GLAND, SPLEEN, TESTIS AND
CC THYMUS.
CC -!- DEVELOPMENTAL STAGE: IN BRAIN, EXPRESSION REACHES MAXIMUM LEVELS
CC AT DAY 12 OF THE EMBRYONIC STAGE. KEEPS ALMOST SIMILAR LEVELS
CC DURING MOUSE DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Y11342; CAA72181.2; -.
DR EMBL; Y11343; CAB95031.1; -.
DR EMBL; Y11344; CAB95031.1; JOINED.
DR EMBL; Y11345; CAB95031.1; JOINED.
DR EMBL; Y11346; CAB95031.1; JOINED.
DR MGD; MGI:1341828; Siat7c.
DR InterPro; IPR001675; Glyco.trans_29.
DR Pfam; PF00777; Glyco.transf.29; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 9 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 29 305 LUMENAL, CATALYTIC (POTENTIAL).
FT DISULFID 80 229 BY SIMILARITY.
FT CARBOHYD 239 239 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 305 AA; 35414 MW; 63C7498619F6A3F CRC64;

Alignment Scores:
Pred. No.: 6,02e-78 Length: 305
Score: 917.00 Matches: 169
Percent Similarity: 89.90% Conservative: 18
Best Local Similarity: 81.25% Mismatches: 21

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 12:38:03 ; Search time 5026 Seconds
(without alignments)
10532.642 Million cell updates/sec

Title: US-09-714-936-218

Perfect score: 1294

Sequence: 1 ccgaattccgggtcgagc.....ttttcttccttttttttg 1294

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.in.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	759	58.7	3207	9	AK091215	AK091215 Homo sapi	
2	743.4	57.4	1106	6	AX376468	AX376468 Sequence	
3	743.4	57.4	1122	6	AX195188	AX195188 Sequence	
4	643.6	49.7	976	9	HSAS07291	HSAS07291 Homo sapi	
5	508	39.3	104714	9	AC105279	AC105279 Homo sapi	
6	508	39.3	166441	9	AC099672	AC099672 Homo sapi	
c	7	508	39.3	191978	9	AC098655	AC098655 Homo sapi
8	506.4	39.1	156585	2	AC012109	AC012109 Homo sapi	
9	449.4	34.7	3903	10	MMY11342	Y11342 Mus musculu	
10	438.2	33.9	2956	10	PATA265	L29554 Rattus norv	
c	11	411.8	31.8	99680	9	AC099769	AC099769 Homo sapi
12	411.8	31.8	127234	2	AL139155	AL139155 Homo sapi	
13	411.8	31.8	197372	9	AC103592	AC103592 Homo sapi	
14	298.2	23.0	1489	10	MMY11345	Y11345 Mus musculu	
c	15	298.2	23.0	254933	2	AC133943	Mus muscu
16	288.6	22.3	278638	2	AC097068	AC097068 Rattus no	
17	226.6	17.5	243529	2	AC113771	AC113771 Rattus no	
c	18	198.4	15.3	127234	2	AL139155	Homo sapi
c	19	198.4	15.3	129305	2	AL355983	Homo sapi
c	20	160.6	12.4	122685	9	AC094023	Homo sapi
21	160.6	12.4	148113	2	AL391624	Homo sapi	
22	160.6	12.4	172239	2	AC079307	Homo sapi	
23	160.6	12.4	182529	9	AC092813	Homo sapi	
24	160.6	12.4	202402	2	AC012517	Homo sapi	
25	153.2	11.8	3559	5	FRU17466	Y17466 Fugu rubrip	
26	134.4	10.4	1541	10	MMY11344	Mus musculu	
c	27	134.4	10.4	123456	2	AC103637	Mus muscu
c	28	134.4	10.4	171593	10	AC122052	Mus muscu
c	29	134.4	10.4	226725	2	AC125762	Rattus no
c	30	134.4	10.4	234331	2	AC116267	Rattus no
c	31	134.4	10.4	253765	2	AC133612	Rattus no
c	32	132.8	10.3	246381	2	AC132545	Rattus no
33	111	8.6	124954	5	AL954371	AL954371 Zebrafish	
34	109.4	8.5	167752	2	BX005450	Danio rer	
35	109.4	8.5	169794	2	BX248509	Danio rer	
36	98	7.6	906	6	AX068255	Sequence	
37	98	7.6	1004	9	AB035172	Homo sapi	
38	98	7.6	1416	6	AX040084	Sequence	
39	98	7.6	1556	6	AR306546	Sequence	
40	98	7.6	1556	6	AX061620	Sequence	
41	98	7.6	1669	9	BC036705	Homo sapi	
42	98	7.6	1687	6	AX068265	Sequence	
43	96.4	7.4	1276	9	HSAS21734	Homo sapi	
44	96.4	7.4	1662	9	AK000600	Homo sapi	
45	95.2	7.4	1685	10	AB030836	Mus sapi	

ALIGNMENTS

RESULT 1
AK091215
LOCUS
DEFINITION Homo sapiens cDNA FLJ33896 fis, clone CTONG2007959, highly similar
to ALPHA-N-ACETYLGLACTOSAMINIDE ALPHA-2,6-SIALYLTRANSFERASE (EC
2.4.99.-).
ACCESSION AK091215
VERSION AK091215.1 GI:21749532
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshina, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3207)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .3207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2007959"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="cloning vector: pME18SFL3"
149. .1066
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC03611.1"
/db_xref="GI:21749533"
/translation="MACILARKSVIAVFIAFLFLVLRVNEVFPILLNCFQPG TWKIFSYTPRRLRHYGVINVKTOELQDLCLCAIVSNQMGYQKVGNEIDRSS CWRNNAFTYEDYGRMTIRVVSHTSVPLLIKPNDFYFKEANTIIYIWIQGFPR MKRDQNGIYNMIAKTVGIPNAQIYVTEKRMSCVDGVEKTKGDRVQSGYLSITG WFTFLAMDACYGIHVYGLINDYCTEGYKRPVHYVEQGRDECDXEFLHBPYGG HRFIEKVFPAKAKHRIIFTHPNLTLS"
BASE COUNT 970 a 630 c 637 g 970 t
ORIGIN
Query Match 58.7%; Score 759; DB 9; Length 3207;
Best Local Similarity 99.3%; Pred. No. 7.3e-197;
Matches 762; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 28 CCGGGTCCCTTATTGGATCTCGGGGAATGTGGCTGGAGAGTCTGCCGTGTACC 87
DB 11 CCGGGTCCCTTATTGGATCTCGGGGAATGTGGCTGGAGAGTCTGCCGTGTACC 70
QY 88 AGCTCTCAGCTGCCCGCAGGACTGCCCTGACCCAGGCGCGCGTCTCGTGGCAG 147
DB 71 ASCTCCAGCTGCCCGCAGGACTGCCCTGACCCAGGCGCGCGTCTCGTGGCAG 130
QY 148 GAGGGCGGGAGGCGCCATGCCCTGCATCCTGAAGAGAAAGTCTGTGATGTGTGAGC 207
DB 131 GAGGGCGGGAGGCGCCATGCCCTGCATCCTGAAGAGAAAGTCTGTGATGTGTGAGC 190
QY 208 TTCATACACGCTTCCTTTTCTCTGCTGGTGTGGTCTTTGTAATGAATGATTTCCCA 267
DB 191 TTCATACACGCTTCCTTTTCTCTGCTGGTGTGGTCTTTGTAATGAATGATTTCCCA 250
QY 268 TTGCTACTAACTGCTTTGGACAACCTGGTACAAAGTGGATACCATTCCTCTACACATAC 327
DB 251 TTGCTACTAACTGCTTTGGACAACCTGGTACAAAGTGGATACCATTCCTCTACACATAC 310
QY 328 AGGCGGCCCTTCGAACACTACATGGATACATAAATGTGAAGACACAGAGCCTTTGCAA 387
DB 311 AGGCGGCCCTTCGAACACTACATGGATACATAAATGTGAAGACACAGAGCCTTTGCAA 370

QY 388 CTGGACTGTGACCTTTGTGCCATAGTGTCAAACTCAGGTCCAGATGCTTGGCCAGAGGTG 447
DB 371 CTGGACTGTGACCTTTGTGCCATAGTGTCAAACTCAGGTCCAGATGCTTGGCCAGAGGTG 430
QY 448 GGAATAGATAGATCGATCCTCTCGATTTGGAGAATGAACAATGCCCCCAACCAAGGT 507
DB 431 GGAATAGATAGATCGATCCTCTCGATTTGGAGAATGAACAATGCCCCCAACCAAGGT 490
QY 508 TATGAGAGAGATGTCGGCGCATGACCATGATTCAGTTCGTGCCATACCAGCGTTCCCT 567
DB 491 TATGAGAGAGATGTCGGCGCATGACCATGATTCAGTTCGTGCCATACCAGCGTTCCCT 550
QY 568 CTTTGTCTAAAAAACCCTGATTTATTTTTCAGGAAGCGAATACTACTACTATTGTGTATT 627
DB 551 CTTTGTCTAAAAAACCCTGATTTATTTTTCAGGAAGCGAATACTACTACTATTGTATT 610
QY 628 TGGGACCTTCCCAATATGAGGAAGAATGCAATGGCATCGTTTACAACTGTGAAA 687
DB 611 TGGGACCTTCCCAATATGAGGAAGAATGCAATGGCATCGTTTACAACTGTGAAA 670
QY 688 AAGACAGTTGGTATCTATCCGAATGCCAAATATACGTGACACAGAGAAGCGCATGAGT 747
DB 671 AAGACAGTTGGTATCTATCCGAATGCCAAATATACGTGACACAGAGAAGCGCATGAGT 730
QY 748 TACTGTGATGAGTTTATTAAGAGAAACTGGGAAGACAGGGGGCA 794
DB 731 TACTGTGATGAGTTTATTAAGAGAAACTGGGAAGACAGAGTCCA 777
RESULT 2
AX376468
LOCUS
DEFINITION Sequence 535 from Patent WO0168848.
ACCESSION AX376468
VERSION AX376468.1 GI:19170623
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J., Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding the same
JOURNAL Patent: WO 0168848-A 535 20-SEP-2001; Genentech, Inc. (US)
FEATURES
source
1. .1106
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 325 a 226 c 260 g 295 t
ORIGIN
Query Match 57.4%; Score 743.4; DB 5; Length 1106;
Best Local Similarity 99.9%; Pred. No. 1.3e-192;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 44 TGGATCTCGGGGAATGTGGCTGGAGAGTCTCGCTGGTGTACAGCCTCCAGCCTGGCC 103
DB 1 TGGATCTCGGGGAATGTGGCTGGAGAGTCTCGCTGGTGTACAGCCTCCAGCCTGGCC 60
QY 104 CCAGGACTCCCTGACCCAGGCGCGCCGCTGCTCGTGGCAGGAGGCGCGGAGCG 163
DB 61 CCAGGACTCCCTGACCCAGGCGCGCCGCTGCTCGTGGCAGGAGGCGCGGAGCG 120
QY 164 CCATGGCTGCATCTGAAAGAAAGTCTGTGATTCGTGTGAGCTTCATAGCAGGTCC 223
DB 121 CCATGGCTGCATCTGAAAGAAAGTCTGTGATTCGTGTGAGCTTCATAGCAGGTCC 180
QY 224 TTTTCTGCTGTTGTGGTCTTTGTAATGAAGTGAATTTCCCATTCGTACTAACTGCT 283

FEATURES

source

Location/Qualifiers

1..976
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/cell_line="HI-29"
1..976
/gene="ST6GALNAC III"
22..939
/gene="ST6GALNAC III"
/EC_number="2.4.99.7"
/codon_start=1
/product="alpha 2,6-sialyltransferase"
/protein_id="CAD45371.1"
/db_xref="GI:22799094"
/translation="MACILKRKSVIAVSFIAAFLVLLVRLVNEVNFLLNCRGQPG
TKWIFSYTRRPLRTHYGVNVKQDLCDAIVNSQVMQVQKQVNEIDRS
CIWRNNAPTRGVEDVGRMTIRVSTSVPLLLKNPDYFFKEANTTYIYIWPFRN
MKDNGIYVWMLKTVIYVNAIYVTEKMSYCDGVFKKTKGDKRVQSGSYLSTG
WTFLLAMDACIGHYVGMINDTYCTGIRKVPYHYEGRDECDYFLHHPAYIG
HRFIEKVFVAKWAKRRIIFHPNWTLS"
BASE COUNT 279 a 202 c 232 g 263 t
ORIGIN
Query Match 49.7%; Score 643.6; DB 9; Length 976;
Best Local Similarity 99.4%; Pred. No. 3.3e-165;
Matches 646; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 145 CAGAGGGCGGCGGAGCCGATGGCTGCATCCTGAAGAAAGTCTGTGATGCTG 204
Db 1 CAGAGGGCGGCGGAGGCCATGGCTGCATCCTGAAGAAAGTCTGTGATGCTG 60
QY 205 AGCTTCATAGCAGCGTTCCTTTTCCCTGCTGTTGCGCTCTGTAATGAGTGAATTC 264
Db 61 AGCTTCATAGCAGCGTTCCTTTTCCCTGCTGTTGCGCTCTGTAATGAGTGAATTC 120
QY 265 CCATTGCTACTAAATGCTTTGGACAACCTGGTACAAAGTGGATACCATCTCTACACA 324
Db 121 CCATTGCTACTAAATGCTTTGGACAACCTGGTACAAAGTGGATACCATCTCTACACA 180
QY 325 TACAGGGGGCCCTTCGAACTCATATGGATACATAATATGACAGACAGAGCCCTTG 384
Db 181 TACAGGGGGCCCTTCGAACTCATATGGATACATAATATGACAGACAGAGCCCTTG 240
QY 385 CAACCTGACTGTGACCTTGTGCCATAGTGTCAAACTCAGTGCAGATGGTGCCAGAAG 444
Db 241 CAACCTGACTGTGACCTTGTGCCATAGTGTCAAACTCAGTGCAGATGGTGCCAGAAG 300
QY 445 GTGGAAATGAGATAGATCGATCCCTCCTGCATTTGGAGAAATGAACAATGCCCCACCAA 504
Db 301 GTGGAAATGAGATAGATCGATCCCTCCTGCATTTGGAGAAATGAACAATGCCCCACCAA 360
QY 505 GGTATGACAGAGAGTGTGGCCCGCATGACCATGATTCGAGTGTGTCCTCCATACAGCGTT 564
Db 361 GGTATGACAGAGAGTGTGGCCCGCATGACCATGATTCGAGTGTGTCCTCCATACAGCGTT 420
QY 565 CCTCTTTGCTAAAAACCTGATTTATTTTTCAGGAGCGGAATGCTACTATTGTTGTT 624
Db 421 CCTCTTTGCTAAAAACCTGATTTATTTTTCAGGAGCGGAATGCTACTATTGTTGTT 480
QY 625 ATTTGGGGACCTTTCGCAATATGAGGAAGATGGCAATGGCATCGCTTTACACATGTTG 684
Db 481 ATTTGGGGACCTTTCGCAATATGAGGAAGATGGCAATGGCATCGCTTTACACATGTTG 540
QY 685 AAAAAGACAGTGTGGTATCTATCCGAATGCCAATATAGTGCATCCACAGACAGAGCCGATG 744
Db 541 AAAAAGACAGTGTGGTATCTATCCGAATGCCAATATAGTGCATCCACAGACAGAGCCGATG 600
QY 745 AGTTACTGTGATGGAGTTTAAAGAGAACTGGGAAGGACAGAGGGGCA 794
Db 601 AGTTACTGTGATGGAGTTTAAAGAGAACTGGGAAGGACAGAGTCCA 650

RESULT 5

AC105279

LOCUS

DEFINITION

AC105279 AL355804

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

104714 bp DNA linear PRI 23-APR-2003

Homo sapiens chromosome 1 clone RP5-1153M13, complete sequence.

AC105279 AL355804

AC105279.3 GI:30061454

HTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 104714)

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,

Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and

Haugen,E.D.

Direct Submission

Unpublished

2 (bases 1 to 104714)

Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.

Direct Submission

Submitted (28-DEC-2001) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

3 (bases 1 to 104714)

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,

Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.

Direct Submission

Submitted (23-MAR-2002) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

4 (bases 1 to 104714)

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,

Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and

Haugen,E.D.

Direct Submission

Submitted (23-APR-2003) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

On Apr 23, 2003 this sequence version replaced gi:19697477.

----- Genome Center

Center: University of Washington Genome Center

Center Code: UWGC

Web site: http://www.genome.washington.edu

Contact: uwgchgsu.washington.edu

Drafting Center: SC

----- Project Information

Center project name: chr-1

Center clone name: RP5-1153M13 (sc0860)

----- Summary Statistics

Sequencing vector: plasmid; 4% of reads

Chemistry: Dye-terminator ET; 56% of reads

Chemistry: Dye-terminator Big Dye; 33% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 104702 bases at least Q40

Consensus quality: 104713 bases at least Q30

Consensus quality: 104714 bases at least Q20

Insert size: 104714; sum-of-contigs

Quality coverage: 7.3x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-293P10 (UWGC:sc0368) AC098655, 104714-bp overlap

3': RP11-286H10 (UWGC:sc0344) AC099672, 96047-bp overlap

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:

[illegible]

Chemistry: Dye-terminator ET; 94% of reads
 Chemistry: Dye-terminator Big Dye; 6% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 191831 bases at least Q40
 Consensus quality: 191959 bases at least Q30
 Consensus quality: 191976 bases at least Q20
 Insert size: 191978; sum-of-contigs
 Quality coverage: 8.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP5-829D5 At138797
 3': RP5-963M5 (UWGC:sc0227) AC099769, 72727-bp overlap

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgIII		BcoRI		HindIII	
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
7611	7721	8696	8776	1351	1348
2067	2062	6	<800	6382	6567
6781	6952	5597	5410	512	<800
1060	1046	1427	1389	449	<800
298	<800	74	<800	1182	1162
1589	1636	906	897	2130	2130
11227	11219	3697	3727	165	<800
278	<800	1001	979	349	<800
5922	5658	10193	10026	1041	1013
550	<800	223	<800	365	<800
724	<800	1112	1093	598	<800
7256	6952	339	<800	1387	1348
2080	2231	3915	4163	2312	2333

3432	3369	3414	3370	4861	4963
741	<800	5158	5047	4230	4250
4142	4283	3809	3727	1367	1348
291	<800	1435	1490	5341	5276
3587	3673	489	<800	1615	1607
5508	5419	7141	7077	5055	4963
731	<800	4563	4856	717	<800
6894	6952	1946	1926	5845	5875
5282	5419	3858	3727	10937	10818
2262	2231	5381	5410	5112	4963
1478	1430	443	<800	3656	3692
545	<800	2930	2885	13462	13440
3254	3369	5480	5410	447	<800
78	<800	448	<800	4106	4250
3192	3059	2141	2079	1024	1013
3192	3059	6411	6307	1626	1607
522	<800	2574	2661	16171	16269
5580	5658	7245	7077	138	<800
902	912	2035	2079	171	<800
3475	3369	1255	1264	7820	7857
6779	6952	4452	4388	757	<800
2609	2559	2587	2661	1990	1897
6324	6040	2663	2661	2185	2130
2808	2755	1086	1062	1717	1706
617	<800	86	<800	1454	1461
3253	3369	2087	2079	4370	4250
4262	4283	7747	7601	1631	1607
674	<800	548	<800	3421	3450
50	<800	6363	6307	9	<800
5659	5658	69	<800	979	1013
3884	3673	1554	1613	177	<800
1045	1046	1774	1726	7	<800
1222	1229	3509	3471	1848	1897
993	1046	1674	1613	6207	6310
1570	1636	15635	15903	2599	2652
3011	3059	2149	2261	1346	1348

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 156385)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhvalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Hagan,L., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczyk,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McClurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrel,A., Vassiliev,H., Vo,A., Wheeler,J., WurX., Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Feb 28, 2000 this sequence version replaced gi:6088018. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RN/RepeatMasker.html ----- Genome Center Center : Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L1202 Center clone name: 45_M_21 ----- Summary Statistics Sequencing vector: M13; M77815; 100% of reads Chemistry: Dye-primer-amersham; 5% of reads Chemistry: Dye-terminator Big Dye; 95% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 113936 bases at least Q40 Consensus quality: 133319 bases at least Q30 Consensus quality: 146386 bases at least Q20 Insert size: 160000; agarose-fp Insert size: 155785; sum-of-contigs Quality coverage: 3.1 in Q20 bases; agarose-fp Quality coverage: 3.2 in Q20 base. * NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * * 1 2316: contig of 2316 bp in length * 2317 2416: gap of 100 bp * 2417 12562: contig of 10146 bp in length * 12563 12662: gap of 100 bp * 12663 23766: contig of 11104 bp in length * 23767 23866: gap of 100 bp * 23867 35469: contig of 11603 bp in length * 35470 35569: gap of 100 bp * 35570 48475: contig of 12806 bp in length * 48476 48575: gap of 100 bp * 48576 67870: contig of 19295 bp in length * 67871 67970: gap of 100 bp * 67971 90884: contig of 22914 bp in length * 90885 90984: gap of 100 bp * 90985 112852: contig of 21868 bp in length * 112853 112952: gap of 100 bp * 112953 156585: contig of 43633 bp in length. * * Location/Qualifiers * 1..156385 * /organism="Homo sapiens" * /mol_type="genomic DNA" * /db_xref="taxon:9606"
FEATURES	source

VERSION	Y11342.1	GI:4894174
KEYWORDS	GalNAc alpha-2, 6-sialyltransferase III.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	Lee, Y.C., Kaufmann, M., Kitazume-Kawaguchi, S., Kono, M., Takashima, S., Kurosawa, N., Liu, H., Pircher, H. and Tsuji, S.	
TITLE	Molecular cloning and functional expression of two members of mouse Neuacalpha2,3Galbeta1,3GalNAc galNAcalpha2,6-sialyltransferase family, ST6GalNAc III and IV	
JOURNAL	J. Biol. Chem. 274 (17), 11958-11967 (1999)	
MEDLINE	99223522	
PUBMED	10207017	
REFERENCE	2 (bases 1 to 3903)	
AUTHORS	Tsuji, S.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-FEB-1997) S. Tsuji, The Institute of Physical & Chemical, Research (RIKEN), Glyco Molecular Biology, Frontier Research Program, Wako, Saitama 351-01, JAPAN	
FEATURES	Location/Qualifiers	
source	1..3903	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="ICR"	
	/db_xref="taxon:10090"	
	/tissue_type="brain"	
	/dev_stage="adult"	
gene	1..3903	
	/gene="ST6GalNAcIII"	
CDS	72..989	
	/gene="ST6GalNAcIII"	
	/codon_start=1	
	/product="GalNAc alpha-2,6-sialyltransferase"	
	/protein_id="CAA72181.2"	
	/db_xref="GI:10241806"	
	/db_xref="SPTREMBL:Q9MWU2"	
	/translation="MACILKRPVLVWGFIALCILILAMLVNDATFPLLNCFCQPK TWIPLPTFPQPLRTHYGVINVRQEPLOLNCHIAIVSNQMGVKGVEEDIDHAS CILWRMNNATPGFEEDYGMATVVRVSHTSYPLLLKNPDYEFKEASRIYVINGPFTN MKRDNGIVYMLKHYVDAYPDQAIVYITEQMTHCDDRVFKDETGRDVQSGSYLSTG WTFILAMDACYS:HVYGMINETYCTKGIRKVPYHYHGEKCEKNVLLHEHPAYGG HRFITEKVPFAKWKHRIYVTHPNWLS"	
BASE COUNT	1059 a 923 c 816 g 1105 t	
ORIGIN		
Query Match	34.7%; Score 449.4; DB 10; Length 3903;	
Best Local Similarity	81.7%; Pred. No. 7e-112; Mismatches 116; Indels 0; Gaps 0;	
Matches 519; Conservative	0;	
Qy	160 AGCGCCATGGCGTCATCCTCGAAGAGAAAGTCTGTGATGCTGTGAGCTTCATAGCAGC 219	
Db	66 AGCGCCATGGCGTCATCCTCAAGAGAAAGCCGTCGTCGTGTGTGAGCTTCATAGCCCTG 125	
Qy	220 TTCCTTTTCTCTGGTTCGCTCTGTGAATGAAGTGAATTTCCCATTTGCTACTAATC 279	
Db	126 TGCATCTTCTCTAGCATGCGCCTGTGCAATGATGCGACTTTCCTTGTCTCTGAAC 185	
Qy	280 TGCTTTGGACACCTGGGTACAAATGGATACCATCTCTCTACACATACAGCGGGCCCTT 339	
Db	186 TGCTTTGGACAGCCTAAGACCAATGGATCCCTTTTGCCCTACACATTCAGGCAGCTCTT 245	
Qy	340 CGAACTACTATTGGATACATTAATGTGAAGACACAAAGAGCCTTTGCAACTGGAGTGCAC 399	
Db	246 CGAACTCACTATGGATACATAAATGTGAGGATCCCAAGAGCCTTTGCAACTGTAAGTGAAC 305	
Qy	400 CTTTGTGCCATAGTCTCAACTGAGTCAAGTGGTTGGCCCAAGAGGTGGGAATCAGATA 459	
Db	306 CACTGTGCCATCGTGTCAAAATCAGGTCAAAATGGTTCGGCAGAGAGTGGGGGAAGAGATA 365	
Qy	460 GATCGATCCTCTCGATTGAGGAATGAACATGCCCCCACCAGAAAGGTATGAAGAAGAT 519	

[illegible]


```

Db      366 GACCATGCAFCCTGTCATCTGGAGAAATGAACACACGCCGCCCAACCAAGGCGCTTIGAGGAAGAT 425
QY      520 GTGCGCGCATGACCATGATTCGAGTTGTCTGCCATACACAGCGTTCTCTTTTGTCTATAAA 579
Db      426 GTGCGGTACATGACATGATGTCGAGTTGTCTGACATACCATGTCCTCTCTGCTGA 485
QY      580 AACCTGATATATTTTCAAGAAAGCAAGCAATACCTACTATTTATTTGTTATTTGGGACCTTTC 639
Db      486 AATCTTGACTATTTTCAAGAAAGCAAGCAAGCAATTTATTTATCTGGGCGCTTTC 545
QY      640 CGCAATATGAGGAAGATGGAATGGCATGCTTTTACACATGTTGAAAAGACAGTTGTT 699
Db      546 CGCACATGAGGAAGATGGAATGGCATGCTTTTACACATGTTGAAAAGACAGTTGTT 605
QY      700 ATCTATCCGAATGCCCAATATAGCTGACACACAGAGAGCGCATGAGTTACTTGTATGGA 759
Db      606 GCCTATCCAGATCGCGAGATCTACGTGACACAGAGACAGATGACTCACTGCGATAGA 665
QY      760 GTTTTAAAGAAAGAACTGGGAAGACAGAGGGGCA 794
Db      666 GTGTTTAAAGATGAGACTGGAAGGACAGAGATCCA 700

RESULT 10
RATA26S
LOCUS      Rattus norvegicus alpha 2,6-sialyltransferase mRNA, complete cds.
DEFINITION
ACCESSION L29554
VERSION    L29554.1 GI:1008902
KEYWORDS   alpha 2,6 sialyltransferase; sialyltransferase.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
REFERENCE  1 (bases 1 to 2956)
AUTHORS    Sjoberg,E.R., Kitagawa,H., Glushka,J., van Halbeek,H. and Paulson,J.C.
TITLE      Molecular cloning of a developmentally regulated
           N-acetylgalactosamine alpha2,6-sialyltransferase specific for
           sialylated glycoconjugates
JOURNAL    J. Biol. Chem. 271 (13), 7450-7459 (1996)
MEDLINE    96205924
PUBMED     8631773
COMMENT    Original source text: Rattus norvegicus (strain Sprague-Dawley)
           adult brain cDNA to mRNA.
           Location/Qualifiers
           1. .2956
           /organism="Rattus norvegicus"
           /mol_type="mRNA"
           /strain="Sprague-Dawley"
           /db_xref="taxon:10116"
           /tissue_type="brain"
           /dev_stage="adult"
           226. .1143
           /codon_start=1
           /product="alpha 2,6-sialyltransferase"
           /protein_id="AAC42086.1"
           /db_xref="GI:1280387"
           /translation="MACILKRKPALAVSFIALCILLLAMRLANDVTFPLLNCFGQPK
           TKWPIASYLRPLQTHYGYINVTQEPFLQNCNCVAVSNQGVQKVEIDRAS
           CIWRNNAPTKGFEDVGMVMTVSVPLLLKNPDYFKEASTYIVYWGPPRN
           MRKNGCIVYMLKTVADYDPAOIYVTTQRTYCDGVFDEKGRVQSGSYLSTG
           WFTPLADACYSIHVYGNINETYCTTEGYKVPYHYVEQKDECNELLYHERAPYGG
           HRFTKKVFAWAKRHLVTFHENVWS"
BASE COUNT      805 a 712 c 665 g 774 t
ORIGIN
Query Match      33.9%; Score 438.2; DB 10; Length 2956;
Best Local Similarity 80.6%; Pred. No. 8.2e-109;
Matches 512; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY      160 AGCGCCATGCGCTCATCTGCTGAAGAGAAAGTCTGTGATGCTGTGAGCTTCATAGCAGCG 219

```

```

Db      220 AGCGCCATGCGCTCATCTGCTGAAGAGAAAGCTTGGGCTTGCAGCTTCATAGCCCTG 279
QY      220 TTCTTTTCTCTCTGCTGCTGCTTGTAAATGAAGTGAATTTCCCATGCTTACCTAAAC 279
Db      280 TGCACTTCTGCTGCTAGCCCTTGGCAATGACGTTCCTTCCCTTTGCTTCCGAC 339
QY      280 TGCTTTGSAACACCTGSPACAAAGTGGATACCATCTCTCTACACATACAGGGGGCCCTT 339
Db      340 TGCTTTGGCAACACCTFAAAACCAATGGATGCCCTTTGCTCTACACATCAGGACGCTCTT 399
QY      340 CGCACTCACTATGATACATTAATGTAAGACACACAGAGCCCTTTGCACTGGAGCTGAC 399
Db      400 CAACCTCACTATGATACATTAATGTAAGACACACAGAGCCCTTTGCACTGGAGCTGAC 459
QY      400 CTTTGTGCCATAGTGTCAAACTCAGGTGAGTGGTGGCCAGAAAGTGGGAAATGAGATA 459
Db      460 CATTTGCCGCTGCTGCGAATGAGCAATGAGTGGTGGAGAGAGTGGGGAGAGATA 519
QY      460 GATCGATCTCTCTGCAATTTGGAGAAATGACATGCCCCCACCACCAAGGTTATGAGAGAT 519
Db      520 GACCGTCCGCTGCTGCAATTTGGAGAAATGAGCAATGAGTGGTGGAGAGAGTTCGAGGAGAT 579
QY      520 GTGCGCGCATGACCATGATTCGAGTTGTGCCATACACAGCGTTCCTCTTTTGTCTAAAA 579
Db      580 GTGCGTACATGACGATGTCGAGTGGTGTACACACACAGCGTCCCTCTCTTGTCTGAAA 639
QY      580 AACCTGATATTTTTCAGGAAGCGAATPACTACTATTTGTGTTATTTGGGACCTTTC 639
Db      640 AACCGGACTATTTTTCAGGAGGCGAGCACAACCATTTACGTCATCTCTGGGCGCTTTC 699
QY      640 CGCAATATGAGAAAGATGCGCAATGGCATCGTTTACACATGTTGAAAAGACAGTTGCT 699
Db      700 CGCAATATGAGAAAGACGCGGATCGTGTACAACATGCTFAAAGAGAGCGGTCGAC 759
QY      700 ATCTATCCGAATGCCCAATATAGTGCACACACAGAGCGCATGACTTACTGTGATGGA 759
Db      760 GCCTACCCAGACGCGCAGATCTAGTGCACACGAGCAGCGGATGAGTACTGGACGGG 819
QY      760 GTTTTAAAGAAAGAACTGGGAAGACAGAGGGCA 794
Db      820 GTGTTTAAAGATGAAACTGGGAAGACAGAGTCCA 854

RESULT 11
AC099769/6
LOCUS      Homo sapiens chromosome 1 clone RP5-963M5, complete sequence.
DEFINITION
ACCESSION AC099769 AL513182
VERSION    AC099769.2 GI:21263323
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 99680)
AUTHORS    Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.
TITLE      Direct Submission
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 99680)
AUTHORS    Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.
TITLE      Direct Submission
JOURNAL    Submitted (20-NOV-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE  3 (bases 1 to 99680)
AUTHORS    Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.
TITLE      Direct Submission
JOURNAL    Submitted (30-MAY-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

```



```
/chromosome="1"
/clone="RP5-963M5"
/clone_lib="RPC1 human PAC library 5"
1..4
/feature="Single subclone region"
4825..5014
/feature="Single subclone region"
34768..34941
/feature="Single subclone region"
80767..80939
/feature="Single subclone region"
33193 a 18620 c 19017 g 28850 t

Query Match 31.8% Score 411.8; DB 9; Length 99680;
Best Local Similarity 99.5%; Pred. No. 1.6e-101;
Matches 413; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 377 AGCCTTGCACCTGAGTGGACCTTTGTGCCATAGTGTCAAACTCAGGTTCAGATGGTTG 436
|||||
Db 90176 AGCCTTGCACCTGAGTGGACCTTTGTGCCATAGTGTCAAACTCAGGTTCAGATGGTTG 90117

QY 437 GCCAGAGGTGGAAATGAGATGATGATCGATCTCTGATTTGGAGATGAACAATGCC 496
|||||
Db 90116 GCCAGAGGTGGAAATGAGATGATGATCGATCTCTGATTTGGAGATGAACAATGCC 90057

QY 497 CCACCAAGGTTATGAGAGATGTCGGCCGATGACCATGATTCGAGTTGTGCCATA 556
|||||
Db 90056 CCACCAAGGTTATGAGAGATGTCGGCCGATGACCATGATTCGAGTTGTGCCATA 89997

QY 557 CCAGCGTTCCCTCTTTTCTAAAAACCTTGATTTTCAAGGAGCGAATACTACTA 616
|||||
Db 89996 CCAGCGTTCCCTCTTTTCTAAAAACCTTGATTTTCAAGGAGCGAATACTACTA 89937

QY 617 TTGTGTTATTTGGGGACCTTCCGCAATATGAGAAAGATGCGAATGCGATCGTTTACA 676
|||||
Db 89936 TTATGTTATTTGGGGACCTTCCGCAATATGAGAAAGATGCGAATGCGATCGTTTACA 89877

QY 677 ACATGTTGAAAAGACAGTTGTTATCTATCCGATGCCAATATACGTGACACAGAGA 736
|||||
Db 89876 ACATGTTGAAAAGACAGTTGTTATCTATCCGATGCCAATATACGTGACACAGAGA 89817

QY 737 AGCCGATGAGTTACTGTGAGTTTAAAGAGAACTGGGAAAGGACAGGGG 791
|||||
Db 89816 AGCCGATGAGTTACTGTGAGTTTAAAGAGAACTGGGAAAGGACAGGGG 89762

FEATURES
source
1..127234
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP4-800D18"
/clone_lib="RPC1-4"
1..2241
/feature="assembly_fragment:00851"
clone_end:SP6
vector_side:left
2342..11262
/feature="assembly_fragment:00081"
11363..54059
/feature="assembly_fragment:00319"
54160..61431
/feature="assembly_fragment:00328"
61532..64263
/feature="assembly_fragment:00383"
64364..71431
/feature="assembly_fragment:00450"
71532..82012
/feature="assembly_fragment:00486"
82113..85208
/feature="assembly_fragment:00507"
85309..88169
/feature="assembly_fragment:00577"

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 119643 bases at least Q40
Consensus quality: 122904 bases at least Q30
Consensus quality: 124553 bases at least Q20
Insert size: 126034; sum-of-contigs
Quality coverage: 3.77x in Q20 bases; sum-of-contigs Quality
coverage: 3.86x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2241: contig of 2241 bp in length
* 2242 2341: gap of 100 bp
* 2342 11262: contig of 8921 bp in length
* 11263 11362: gap of 100 bp
* 11363 54059: contig of 42897 bp in length
* 54060 54159: gap of 100 bp
* 54160 61431: contig of 7272 bp in length
* 61432 61531: gap of 100 bp
* 61532 64263: contig of 2732 bp in length
* 64264 71431: contig of 7088 bp in length
* 71432 71531: gap of 100 bp
* 71532 82012: contig of 10481 bp in length
* 82013 82112: gap of 100 bp
* 82113 85208: contig of 3096 bp in length
* 85209 85308: gap of 100 bp
* 85309 88169: contig of 2861 bp in length
* 88170 88269: gap of 100 bp
* 88270 108304: contig of 20035 bp in length
* 108305 108404: gap of 100 bp
* 108405 116626: contig of 8222 bp in length
* 116627 116726: gap of 100 bp
* 116727 119049: contig of 2323 bp in length
* 119050 127234: gap of 100 bp
* 127234: contig of 8085 bp in length.
Location/Qualifiers
1..127234
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP4-800D18"
/clone_lib="RPC1-4"
1..2241
/feature="assembly_fragment:00851"
clone_end:SP6
vector_side:left
2342..11262
/feature="assembly_fragment:00081"
11363..54059
/feature="assembly_fragment:00319"
54160..61431
/feature="assembly_fragment:00328"
61532..64263
/feature="assembly_fragment:00383"
64364..71431
/feature="assembly_fragment:00450"
71532..82012
/feature="assembly_fragment:00486"
82113..85208
/feature="assembly_fragment:00507"
85309..88169
/feature="assembly_fragment:00577"
```

[illegible]

QY	497	GCACCAAGGATGGGAATAGAGATCGCTCCCTGCTTTCGAGHGTGACCATATGCC	478
Dd	121762	GCCACAAGGTGGGAATAGAGATGATCCTCTCGCATTTGAGAATGAACAATGCC	121821
QY	497	CCACCAAAGGTTTAGAAGAAGATGTCGGCCGATGACCATGATTCGAGTTGTGTCCCATA	556
Dd	121822	CACCAAAGGTTTAGAAGAAGATGTCGGCCGATGACCATGATTCGAGTTGTGTCCCATA	121881
QY	557	CCAGCGTCCCTCTTTGCTAAAAAACCCGTGATTATTTTTTCAAGAAGCGCAATACTACTA	616
Dd	121882	CNAGCGTTCCTCTTTGCTTAAAAAACCCGTGATTATTTTTTCAAGAAGCGCAATACTACTA	121941
QY	617	TTTGTTGTTATTGGGACCTCTCCGCAATATGAGGAAGATGGCAATGGCCTGTTACA	676

677	ACATGTTGAAAAGACAGATGGTGTGTCATCCGAATCCCAATATATCGTGACCACAGAGA	736	
QY			
DB	122002	ACATGTTGAAAAGACAGATGGTGTCTATCCGAATCCCAATATATCGTGACCACAGAGA	122061
QY	737	AGCGCATCAGTTACTGTGATGGAGTTTTTAAGAAGAACTGGGAAGGCACAGGGG	791

122002	AC006416GSI	AC101610GSI	111111	ACGAGGAGGAC	1036636GAC	AGAGG	122110
RESULT 13							
AC103592							
LOCUS							
AC103592 197372 bp DNA linear PRI 10-AUG-2002							
DEFINITION Homo sapiens chromosome 1 clone RP11-335F14, complete sequence.							
ACCESSION							
AC103592							
VERSION							
AC103592.2 GI:22203287							
KEYWORDS							
HTG.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 197372)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 197372)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (29-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 197372)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
Direct Submission
Unpublished

REFERENCE 1
AUTHORS Lee,Y.C., Kaufmann,M., Kitazume-Kawaguchi,S., Kono,M., Takashima,S., Kurosawa,N., Liu,H., Pircher,H. and Tsuji,S.
TITLE Molecular cloning and functional expression of two members of mouse Neuocalpha2,3galactal,3galnac galnacalpa2,6-sialyltransferase family, ST6Galnac III and IV
JOURNAL J. Biol. Chem. 274 (17), 11958-11967 (1999)
MEDLINE 99223522
PUBMED 10207017
REFERENCE 2 (bases 1 to 1489)
AUTHORS Tsuji,S.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-1997) S. Tsuji, The Institute of Physical & Chemical, Research (RIKEN), Glyco Molecular Biology, Frontier Research Program, Wako, Saitama 351-01, JAPAN
FEATURES
SOURCE 1..1489
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/cell_line="NIH3T3"
837..1246
/gene="ST6GalnacIII"
837..1246
/gene="ST6GalnacIII"
/number=3
/usedin=Y11343:III_CDS
/usedin=Y11343:III_RNA
BASE COUNT 434 a 321 c 306 g 426 t 2 others
ORIGIN
Query Match 23.08; Score 298.2; DB 10; Length 1489;
Best Local Similarity 82.4%; Pred. No. 2.1e-70;
Matches 342; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 377 AGCCTTTCACCTGAGTGTGACCTTTGGCCATAGTGTCAAACTCAGGTGAGTGGTGG 436
DB 835 AGCCTTTCACCTGAGTGTGACCTTTGGCCATAGTGTCAAACTCAGGTGAGTGGTGG 894
QY 437 GCACAGAGTGGGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 496
DB 895 GGCAGAGTGGGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 954
QY 497 CCACCAAGGTTATGAGAGAGATGCGGCCATGATGATGATGATGATGATGATGATG 556
DB 955 CACCAAGGCTTTGAGAGAGATGCGGCCATGATGATGATGATGATGATGATGATG 1014
QY 557 CCAGCGTCTCTCTTTTCT 616
DB 1015 CCAGTGTCT 1074
QY 617 TTGTGTATTTGGGGACCTTCCGCAATATGAGGAAGATGCAATGCGTCTCTCTCTCT 676
DB 1075 TTATGTATTTGGGGACCTTCCGCAATATGAGGAAGATGCAATGCGTCTCTCTCTCT 1134
QY 677 ACATGTGTAAGAGACAGTGGTATCTATCCGAATGCCAATATACGTGACCAAGAGA 736
DB 1135 ACATGTGTAAGAGACAGTGGTATCTATCCGAATGCCAATATACGTGACCAAGAGA 1194
QY 737 AGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 791
DB 1195 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
RESULT 15
AC133943/c
LOCUS AC133943 254933 bp DNA linear HTG 21-SEP-2002
DEFINITION Mus musculus chromosome UNK clone RP24-104D12, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
AC133943
AC133943.1
VERSION AC133943.1 GT:23238072
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 254933)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone unpublished
JOURNAL 2 (bases 1 to 254933)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 254933)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BB0104D12
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing method: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 249849 bases at least Q40
Consensus quality: 250816 bases at least Q30
Consensus quality: 251266 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1651: contig of 1651 bp in length
* 1751: gap of unknown length
* 3774: contig of 2023 bp in length
* 3775: gap of unknown length
* 5094: contig of 1220 bp in length
* 5194: gap of unknown length
* 5195: contig of 4423 bp in length
* 9618: gap of unknown length
* 9718: contig of 8439 bp in length
* 18157: gap of unknown length
* 18257: contig of 29739 bp in length
* 47996: gap of unknown length
* 48096: contig of 33795 bp in length
* 81891: gap of unknown length
* 81891: contig of 39164 bp in length
* 121155: gap of unknown length
* 121255: contig of 5517 bp in length
* 176372: gap of unknown length
* 176472: contig of 78462 bp in length.
FEATURES
Location/Qualifiers
1..254933
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-104D12"
1..1651
/note="assembly_name:Contig10"
misc_feature

misc_feature 1752..3774 /note="assembly_name:Contig11"
misc_feature 3875..5094 /note="assembly_name:Contig12"
misc_feature 5195..9617 /note="assembly_name:Contig13"
misc_feature 9718..18156 /note="assembly_name:Contig14"
misc_feature 18257..47995 /note="assembly_name:Contig15"
misc_feature 48096..81890 /note="assembly_name:Contig16"
misc_feature 81991..121154 /note="assembly_name:Contig17"
misc_feature 121255..176371 /note="assembly_name:Contig18"
misc_feature 176472..254933 /note="assembly_name:Contig19"
BASE COUNT 70197 a 56746 c 55884 g 71184 t 922 others
ORIGIN

Query Match 23.0%; Score 298.2; DB 2; Length 254933;
Best Local Similarity 82.4%; Pred. No. 2.6e-70;
Matches 342; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 377 AGCCTTGCACCTGACCTTTGGCCATAGTGTCAAACTCAGGTCAGATGGTTG 436
Db 84149 AGCCTTGCACCTGACCTTTGGCCATAGTGTCAAACTCAGGTCAGATGGTTG 84090
QY 437 GCCAAGAGTGGGAATGAGATAGATCGATCCCTCTGATTTGGAGATGAACAATGCC 496
Db 84089 GCCAAGAGTGGGAAGAGATAGACCATCATCTCTGATCTGGAGATGAACAATGCC 84030
QY 497 CCACCAAGGTATGAAGAAGATGTCGGCGCATGACCATGATTCGAGTTGTGTCCATA 556
Db 84029 CAACCAAGGCTTTGAGGAAGATGTCGGCTACATGACAATGTCGAGTTGTGTACATA 83970
QY 557 CCAGCGTCTCTTTTGTAAACCCCTGATTTATTTTCAAGSAGCCGATACTACTA 616
Db 83969 CCAGTGCCTCTCTGTGTGAAATAATCCTGACTATTTTTCAAAGAGCAGGAGCCA 83910
QY 617 TTTGTGTTATTTGGGACCTTTCCGCAATATGAGAAAGATGGCAATGGCATCGTTTACA 676
Db 83909 TTTATGTTATCTGGGGCCCTTTCCGACATGAGGAGGATGGTATGGCATCGTGTACA 83850
QY 677 ACATGTTGAAAAGACAGTTGGTATCTATCCGAATCCCAATATACGTGACCACAGAGA 736
Db 83849 ACATGTTAAGAAGACAGTTGATGCTATCCAGATCCGAGATCTACGTGACCACAGAGC 83790
QY 737 AGCGCATGAGTTAGTGTGAGTTTAAAGAGGAAACTGGGAAGGACAGGGG 791
Db 83789 AGCAGATGACTCTCGCATAGATGTTTAAAGATGAGACTGGAAAGGACAGGTG 83735

Search completed: September 12, 2003, 14:10:54
Job time : 5035 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:20:14 ; Search time 19.5 Seconds
(without alignments)
6241.288 Million cell updates/sec

Title: US-09-714-936-218
Perfect score: 2383
Sequence: 1 ccggaattccggcgagc.....ttttctctctttttttg 1294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framet-n2p.model -DEV=xlh
-Q/cyn2_1/USPTO.spool/US09714936/runat_10092003_151458_5680/app_query.fasta_1.1479
-DB=SwissProt_41 -QPM=fastan -SUFFIX=n3p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714936 -CGN_1_17 -runat_10092003_151458_5680 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	941	39.5	305	1	SI7C_RAT	Q64686 rattus norv
2	917	38.5	305	1	SI7C_MOUSE	Q9wuv2 mus musculu
3	415	17.4	360	1	SI7D_MOUSE	Q9r2b6 mus musculu
4	382	16.0	302	1	SI7D_HUMAN	Q9h4f1 h alpha-n-a
5	328.5	13.8	336	1	SI7E_HUMAN	Q9bvh7 homo sapien
6	319.5	13.4	336	1	SI7E_MOUSE	Q9qj71 mus musculu
7	163	6.8	413	1	SI7A_CHICK	Q92182 gallus gall
8	146.5	6.1	403	1	SI7A_MOUSE	Q64685 mus musculu
9	142	6.0	406	1	SI7A_HUMAN	P15907 homo sapien
10	138.5	5.8	374	1	SI7A_RAT	Q02734 x cmp-n-ace
11	138.5	5.8	403	1	SI7A_RAT	P13721 rattus norv
12	137.5	5.8	374	1	SI7A_MOUSE	P97325 mus musculu
13	137.5	5.8	375	1	SI7A_HUMAN	Q11203 homo sapien
14	136	5.7	342	1	SI7A_CHICK	Q11200 g cmp-n-ace
15	135.5	5.7	350	1	SI7B_MOUSE	Q11204 m cmp-n-ace
16	135.5	5.7	350	1	SI7B_MOUSE	Q11205 x cmp-n-ace
17	134	5.6	340	1	SI7B_RAT	Q11201 h cmp-n-ace
18	131.5	5.5	350	1	SI7B_HUMAN	Q16842 h cmp-n-ace

19	131	5.5	337	1	SI4A_MOUSE	P54751 m cmp-n-ace
20	124	5.2	380	1	SI8C_MOUSE	Q64689 mus musculu
21	123.5	5.2	343	1	SI4A_PIG	Q02745 s cmp-n-ace
22	119	5.0	359	1	SI49_MOUSE	Q88829 mus musculu
23	118.5	5.0	333	1	SI4C_HUMAN	Q11206 h cmp-n-ace
24	116.5	4.9	404	1	SI7B_CHICK	Q92184 gallus gall
25	114	4.8	380	1	SI8C_HUMAN	Q43173 homo sapien
26	113.5	4.8	355	1	SI8A_MOUSE	Q64687 mus musculu
27	112.5	4.7	362	1	SI49_HUMAN	Q9unp4 homo sapien
28	112	4.7	526	1	SI7A_MOUSE	Q9g239 mus musculu
29	110	4.6	566	1	SI7A_CHICK	Q92183 gallus gall
30	109.5	4.6	373	1	SI7B_MOUSE	P70277 mus musculu
31	108.5	4.6	600	1	SI7A_HUMAN	Q9nsc7 homo sapien
32	106	4.4	356	1	SI8A_HUMAN	Q92185 homo sapien
33	105.5	4.4	374	1	SI7B_HUMAN	Q9uj37 homo sapien
34	103	4.3	375	1	SI8B_MOUSE	Q35696 mus musculu
35	101	4.2	375	1	SI8B_HUMAN	Q92186 homo sapien
36	100	4.2	375	1	SI8B_RAT	Q07977 rattus norv
37	97.5	4.1	465	1	GLGA_SYNP7	Q935y7 synechococc
38	92.5	3.9	759	1	PIGQ_HUMAN	Q9brb3 homo sapien
39	92	3.9	359	1	SI8D_HUMAN	Q92187 homo sapien
40	92	3.9	1877	1	PKC5_MOUSE	Q04592 mus musculu
41	89.5	3.8	712	1	S21F_HUMAN	Q9nyb5 homo sapien
42	89.5	3.8	2282	1	ZAN_RABIT	P57999 oryctolagus
43	87.5	3.7	173	1	CRGA_MOUSE	P04345 mus musculu
44	87.5	3.7	387	1	CYB_SCHPO	P05501 schizosacch
45	87	3.7	216	1	NKGD_HUMAN	P26718 homo sapien

ALIGNMENTS

RESULT 1

SI7C_RAT	SI7C_RAT	STANDARD;	PRT;	305 AA.
AC	Q64686;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Alpha-N-acetyl-galactosaminide alpha-2,6-sialyltransferase III			
DE	(EC 2.4.99.-) (S76GALNAc III) (Sialyltransferase 7C) (STY).			
GN	STAT7C.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RX	MEDLINE=96205924; PubMed=8631773;			
RA	Sjoberg E.R., Kitagawa H., Glushka J., van Halbeek H., Paulson J.C.;			
RT	"Molecular cloning of a developmentally regulated			
RT	N-acetyl-galactosamine alpha2,6-sialyltransferase specific for			
RT	sialylated glycoconjugates.";			
RL	J. Biol. Chem. 271:7450-7459(1996).			
CC	!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminate + alpha-N-			
CC	acetylneuraminyl-2,3-beta-D-galactosyl-1,3-(N-acetyl-D-			
CC	galactosaminyl)-glycoprotein = CMP + alpha-N-acetylneuraminyl-			
CC	2,3-beta-D-galactosyl-(2,6-alpha-N-acetylneuraminyl)-(N-acetyl-D-			
CC	galactosaminyl)-glycoprotein.			
CC	!- PATHWAY: Glycosylation.			
CC	!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).			
CC	!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, NEWBORN BRAIN AND			
CC	KIDNEY AND TO A LESSER EXTENT IN LUNG. NOT FOUND IN LIVER AND			
CC	SKELETAL MUSCLE.			
CC	!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.			

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase III
DE	(EC 2.4.99.-) (S76GALNAC III) (Sialyltransferase 7C) (SYT).
GN	SIA7C.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID:10090;
RN	[1]
RP	SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP	STAGE.
RC	STRAIN=ICR; TISSUE=Brain;
RX	MEDLINE=9923522; PubMed=10207017;
RA	Lee Y.-C., Kaufman M., Kitazume-Kawaguchi S., Kono M., Takashima S.,
RA	Kurosawa N., Liu H., Picher H., Tsuji S.;
RT	"Molecular cloning and functional expression of two members of mouse
RT	Neuracalphat2,3galbetal,3galnac Galnrcalphat2,6-sialyltransferase family,
RT	Sf6galnac III and IV.";
RL	J. Biol. Chem. 274:11958-11967(1999).
CC	-!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF GANGLIOSIDE GDIA FROM
CC	GMBL. TRANSFERS CMP-NEUAC WITH AN ALPHA-2,6-LINKAGE TO GALNAC
CC	RESIDUE ON NEUAC-ALPHA-2,3-GAL-BETA-1,3-GALNAC OF GLYCOPROTEINS
CC	AND GLYCOLIPIDS. S76GALNACIII PREFERS GLYCOLIPIDS TO
CC	GLYCOPROTEINS.
CC	-!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminate + alpha-N-
CC	acetylneuraminyl-2,3-beta-D-galactosyl-1,3-(N-acetyl-D-
CC	galactosaminyl)-glycoprotein = CMP + alpha-N-acetylneuraminyl-
CC	2,3-beta-D-galactosyl-(2,6-alpha-N-acetylneuraminyl)-(N-acetyl-D-
CC	galactosaminyl)-glycoprotein.
CC	-!- PATHWAY: Glycosylation.
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC	-!- TISSUE SPECIFICITY: HIGH EXPRESSION IN BRAIN, LUNG AND HEART AND
CC	TO A LESSER EXTENT IN KIDNEY, MAMMARY GLAND, SPLEEN, TESTIS AND
CC	THYMUS.
CC	-!- DEVELOPMENTAL STAGE: IN BRAIN, EXPRESSION REACHES MAXIMUM LEVELS
CC	AT DAY 12 OF THE EMBRYONIC STAGE. KEEPS ALMOST SIMILAR LEVELS
CC	DURING MOUSE DEVELOPMENT.
CC	-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Y11342; CAB92181.2; -
DR	EMBL; Y11343; CAB95031.1; -
DR	EMBL; Y11344; CAB95031.1; JOINED.
DR	EMBL; Y11345; CAB95031.1; JOINED.
DR	EMBL; Y11346; CAB95031.1; JOINED.
DR	MGI: MGI:1341828; Siat7c.
DR	InterPro; IPR001675; Glyco.trans.29.
DR	Fram; FR00777; Glyco.transf.29; I.
KW	Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW	Signal-anchor; Golgi stack.
FT	DOMAIN 1 8
FT	TRANSMEM 9 28
FT	CYTOPLASMIC (POTENTIAL). SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL)
FT	DOMAIN 29 305
FT	LUMENAL, CATALYTIC (POTENTIAL).
FT	BY SIMILARITY.
FT	DISULFID 80 239
FT	CARBOHYD 239 239
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 301 301
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 305 AA; 35414 MW; 63C7498615BF6A3F CRC64;
Alignment Scores:	
Pred. No.:	6,02e-78
Score:	917.00
Percent Similarity:	89.90%
Best Local Similarity:	81.25%
Mismatches:	21
Length:	305
Matches:	169
Conservative:	18

```

Query Match:      38.48%      Indels:      0
DB:               1          Gaps:        0
US-09-714-936-218 (1-1294) x SI7C_MOUSE (1-305)

QY 166 ATGCGCTGATCTGAAGAAAGTCGTGATTCGTGCTGAGCTCATAGCAGCGTTCCTT 225
DB 1 Metacystis leuLeuysArgLysProValValValSerPheIleAlaLeuCysIle 20
QY 226 TTCCTGCTGTTGTCGGCTCTGTAATCAAGTGAATTTCCCATTCGTACTAACTGCTTT 285
DB 21 LeuLeuLeuAlaMetArgLeuValAsnAspAlaThrPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGCAACCTGGTACAAAGTGAATTCCTCATACATACAGCGCGCCCTTCGAATC 345
DB 41 GlyGlnProLysThrLysTrpIleProLeuProLysThrPheArgGlnProLeuLysArg 60
QY 346 CACTATGGATACATAAATGTGAAGACACAGAGCCCTTGCACACTGGACTGTGACTTGT 405
DB 61 HisTyrGlyThrIleAsnValArgThrGlnGlnProLeuGlnLeuAsnCysAsnHisCys 80
QY 406 GCCATAGTGCAACTCAGGTCAGATGTTGGCCAGAGGTGGGAATGAGATAGATCGA 465
DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyGlnGluIleAspHis 100
QY 466 TCCCTCGCATTTGGAGATGAACAATGCCCCACCAAGGTTATGAGAGAGATGTCGCG 525
DB 101 AlasercysIleTrpArgMetAsnAspAlaProThrLysGlyPheGluGlnAspValGly 120
QY 526 CGCATGACCATGATTCGAGTGTGTGCCATACACAGCGTTCCTCTTTCTTAAACCACT 585
DB 121 TyrMetThrMetValArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGAGAGCGAATACTACTATTGTTGTTATTTGGGACCTTCCGCAAT 645
DB 141 AspTyrPhePheLysGluAlaSerArgThrIleTyrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGGAAGATGGCAATGGCATGTTTACACATGTTGAAAACACAGTGTGCTATCTAT 705
DB 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysThrValAspAlaTyr 180
QY 706 CCGAATGCCCAATATACGTACACACAGAGAGCGCATGAGTACTGTGAGGAGTTT 765
DB 181 ProAspAlaGlnIleTyrValThrThrGluGlnMetThrHisCysAspArgValPhe 200
QY 766 AAGAAGCAACTCGGAGAGCAAGG 789
DB 201 LysAspGluThrGlyLysAspArg 208

RESULT 3
SI7D_MOUSE STANDARD: PRT: 360 AA.
AC Q9R2B6; Q88725; Q9JHP0; Q9QUP9; Q9R2B5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3-N-
DE acetyl-galactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.7)
DE (NeuAc-alpha-2,3-Gal-beta-1,3-GalNAc-alpha-2,6-sialyltransferase)
DE (ST6GalNAc IV) (Sialyltransferase 7D).
GN STAT7D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., ACTIVITY, TISSUE SPECIFICITY, ALTERNATIVE
RP SPLICING, AND DEVELOPMENTAL STAGE.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=99223522; PubMed=10207017;
RA Lee I.-C., Kaufman M., Kitazume-kawaguchi S., Kono M., Takashima S.,
RA Kurosawa N., Liu H., Pircher H., Tsuji S.;
RT "Molecular cloning and functional expression of two members of mouse

```

Alignment Scores:

```

Pred. No.: 8.3e-31 Length: 360
Score: 415.00 Matches: 105
Percent Similarity: 54.91% Conservative: 46
Best Local Similarity: 38.18% Mismatches: 97
Query Match: 17.42% Indels: 27
DB: 1 Gaps: 8

US-09-714-936-218 (1-1294) x SI7D_MOUSE (1-360)
QY 38 CTTATTGGATCGGGGAATGGCGCTGGA---GAGGTCTCGCTCGGTGATACCACTCC 94
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 Leu1eTxlPleCysLeuSerLeu-AlaSerValSerLeuIleSerTlPAspGlnLeuPr 38
QY 95 AGCC-----TCCGCCAGACAGTCCCTCGACCCAGCGCGCGCTGCTGGTGGCAGG 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 38 oAlaPheLeuIleProSer-Thr-Gly---AspSerSerLeuGlnThrAlaLysSerArg 57
QY 149 AGGCGCGGAGCGGACCCATCGCTGATCGTGAAGAGAAAGTGTGATGCTGTGAGCT 208
   :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 spSerMetLysAlaProGlyArgLeuLeuLeuThrLeuLeuLeuThrPheSerA 77
QY 209 TCATACAGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 268
   :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 laValCysValPheLeu-CysCysTrpAlaCys-----LeuProLeu 90
QY 269 TGCTACTAACTGCTTGGACACCTGGTACAAAGTGGATACCACTTCTCTCTCTCTCT 319
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 CysLeuAlaThrCysLeuAspArg-----HisLeuProAlaAlaPro 104
QY 320 ---ACACAPACAGGCGGCGCCCTCGAATCTACTATGATACATAAATGTGAAGACAC 376
   :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ArgSerThrValProGlyProLeuHisPheSer---GlyTyrSerSerValProAspGlyL 124
QY 377 AGCTTTTGAACGTGAC---TGTGACCTTTGTGCGATAGTCTCAACTGCTGAGTGG 433
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 ysProLeuIleArgGlnLeuCysHisSerCysAlaValValSerAsnSerGlyGlnMetL 144
QY 434 TTGGCCAGAGGTGGGAAATGAGATGATGATGATGATGATGATGATGATGATGATGAT 493
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 euGlySerGlyLeuGlyAlaGlnIleAspGlyAlaGlyCysValLeuArgMetAsnGlnA 164
QY 494 CCCCACCAAGGTATGACAGAGATGTCGGCGCATGACCATGATTCGATGTCGTCCTCC 553
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 laProThrValGlyPheGlnGluAspValGlyGlnArgThrLeuArgValIleSerH 184
QY 554 ATACACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 613
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 isThrSerValProLeuLeuLeuArgAsnTyr-SerHisTyrPheGlnHisAlaArgAspT 204
QY 614 CTATTGTGTTATTTGGGACCTTCCGCATATGAGAAAGATGCGCATGCGCATGCTTT 673
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 hrLeuTyrValValTrpGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThr 224
QY 674 ACAACATGTTGAAAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 733
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 yrArgThrLeuLeuGlnLeuThrArgMetTyr-ProGlyLeuGlnValTyrThrPheThrG 244
QY 734 AGAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 luArgMetMetAlaTyrCysaspGlnIlePheGlnAspGluThr-GlyLysAsnArgArgG 264
QY 794 ATGACAGGCGATGCTGATTTCTACACACACTTTT 828
Db 264 lnSerGlySerPheLeu---SerThrGlyTrpPhe 274

RESULT 4
SI7D_HUMAN STANDARD: PRT: 302 AA.
AC Q9H4F1; Q9NWU6; Q9UKU1; Q9ULB9; Q9Y3G3; Q9Y3G4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3-N-

```

```

DE DE acetyl-galactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.7)
DE DE (NeuAc-alpha-2,3-Gal-beta-1,3-GalNAc-alpha-2,6-sialyltransferase)
DE DE (ST6GALNAC IV) (Sialyltransferase 7D) (Sialyltransferase 3C).
CN SIAT7D OR SIAT3C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RN MEDLINE=20517254; PubMed=11062056;
RA Harduin-Lepers A., Stokes D.C., Steelant W.F.A., Samyn-Petit B.,
RA Krzewinski-Recochi M.A., Vallejo-Ruiz V., Zanetta J.P., Auge C.,
RA Delannoy P.;
RT "Cloning, expression and gene organization of a human Neu5Ac alpha
RT 2-3Gal beta 1-3GalNAc alpha 2,6-sialyltransferase: hST6GALNACIV.";
RL Biochem. J. 352:37-48(2000).
RN [2]
RN SEQUENCE FROM N.A.
RN TISSUE=Petal liver;
RC Kim K.-W., Kim K.-S., Do S.-I., Kim C.-H., Lee Y.-C.;
RT "Molecular cloning of NeuAcalpha2,3Galbeta1,3GalNAc alpha2,6-
RT sialyltransferase cDNA from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RN TISSUE=Lung;
RC Yoshida A.;
RT "N-acetylgalactosaminide alpha2,6-sialyltransferase.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Oobayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isoqai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE OF 213-302 FROM N.A.
RA MEDLINE=99299247; PubMed=10369878;
RX Gilley J., Fried M.;
RT "Extensive gene order differences within regions of conserved synteny
RT between the fugu and human genomes: implications for chromosomal
RT evolution and the cloning of disease genes.";
RL Hum. Mol. Genet. 8:1313-1320(1999).
CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF GANGLIOSIDE GD1A FROM
CC GM1B. TRANSFERS CMP-NEUAC WITH AN ALPHA-2,6-LINKAGE TO GALNAc
CC RESIDUE ON NEUAc-ALPHA-2,3-GAL-BETA-1,3-GALNAc OF GLYCOPROTEINS
CC AND GLYCOLIPIDS. PREFERENCES GLYCOPROTEINS TO GLYCOLIPIDS (BY
CC similarity).
CC -1- CATALYTIC ACTIVITY: CMP-N-acetylneuraminyl + alpha-N-
CC acetylneuraminyl-2,3-beta-D-galactosyl-1,3-N-acetyl-D-
CC galactosaminyl-R = CMP + alpha-N-acetylneuraminyl-2,3-beta-D-
CC galactosyl-1,3-(alpha-N-acetylneuraminyl-2,6)-N-acetyl-D-
CC galactosaminyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AJ271734; CAC07404.1; -
CC DR EMBL; AF127142; AAF00102.1; -
CC DR EMBL; AB035172; BAA87034.1; -
CC DR EMBL; AK000600; BAA91281.1; -

```

```

DR EMBL; Y17460; CAB4353.1; -.
DR EMBL; Y17461; CAB4354.1; -.
DR Genew; HGNC:17846; SIAT7D.
DR MIM; 606378; -.
DR InterPro: IPR001675; Glyco_trans_29.
DR Pfam; PF00777; Glyco_transf_29; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 28 302 LUMENAL, CATALYTIC (POTENTIAL).
FT DISULFID 76 225 BY SIMILARITY.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 119 120 S -> T (IN REF. 1).
FT CONFLICT 119 120 ST -> QA (IN REF. 2).
FT CONFLICT 140 140 F -> L (IN REF. 4).
SQ SEQUENCE 302 AA; 34200 MW; 08A4CDC749A6D783 CRC64;

Alignment Scores:
Pred. No.: 9.88e-28 Length: 302
Score: 382.00 Matches: 86
Percent Similarity: 58.57% Conservative: 37
Best Local Similarity: 40.95% Mismatches: 59
Query Match: 16.03% Indels: 28
DB: 1 Gaps: 6

US-09-714-936-218 (1-1294) x SI7D_HUMAN (1-302)
QY 230 TGTGTG-----TTGTGCTGTTGTAATGAAGTGAATTCCTCCATGTCACATA 277
Db 26 CysTrpAlaGlyLeuProLeuCysLeuAla 35
QY 278 ACTGCTTTGGCAACCTGGTACAAAGTGGATACCATCTCTCTACACATACA----- 328
Db 36 ThrCysLeuAsp-----HisHisPheProThrGlySerArgProThr 49
QY 329 ---GGGGCCCTTCGAACTCCTACATGATGATACATAAATGCAAGACACAGAGCTTTGCC 385
Db 50 ValProGlyProLeuHisPheSer---GlyTyrSerValProAspGlyLeuProLeu 69
QY 386 AAGTGGAC---TGTGAOCTTTGTGCCATAGTGTCAAACTCAGGTGAGTGGTGGCCAGA 442
Db 69 AlaArgGluProCysArgSerCysAlaValValSerSerSerGlyGlnMetLeuGlySerG 89
QY 443 AGTGGGAATGAGATGATGATGATCTCTCTGATTTGGAGATGAGATGAGATGAGATGAG 502
Db 89 LeuGlyAlaGluLeuAspSerAlaGluCysValPheArgMetAsnGlnAlaProThrV 109
QY 503 AAGGTTATGAGACATGTCGGCGCATGATGATGATGATGATGATGATGATGATGATGATG 562
Db 109 alGlyPheGluAlaAspValGlyGlnArgSerThrLeuArgValValSerHisThrSerV 129
QY 563 TTCTCTTTTGTCTAAAAACCCGATATTATTTTCAAGAGAGCGAATACTACTATTGTG 622
Db 129 alProLeuLeuLeuArgAspTyrSerHisTyrPheGluAlaAspGspThrLeuTyrM 149
QY 623 TTATTTGGGACCTTCCGCAATATGAGGAAGTGGCAATGCGATCGCTTTACACATGT 682
Db 149 etValTrpGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThrL 169
QY 683 TGAAGAACACATGTTGGTATCTATCCGAATGCCCAATATACGTGACACAGAGAGCGCA 742
Db 169 euLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMetX 189
QY 743 TGAGTTACTGTGATGAGATTTTAAAGAGAAACTGGAAGACAGAGCGGCGCATGCAAGC 802
Db 189 etAlaTyrCysAspGlnIlePheGlnAspGlnThrGlyLeuAsnArgGlnSerGlyS 209
QY 803 GACTGCTGATTCTACACACATTTT 828
Db 209 erPheLeu---SerThrGlyTrpPhe 216

```

```

RESULT 5
SI7E_HUMAN
ID SI7E_HUMAN STANDARD; PRT; 336 AA.
AC Q8BVH7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase V (GSI alpha
DE synthase) (GalNAc alpha-2,6-sialyltransferase V) (ST6GalNAC V)
DE (Sialyltransferase 7E).
DE SIAT7E.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida H., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Nagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eve;
EX MEDLINE=22388257; Pubmed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Involved in the biosynthesis of ganglioside GDla from
CC glycoproteins (By similarity).
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AK056241; BAB71127.1; -.
CC EMBL; BC001201; AA01201.1; -.
CC InterPro: IPR001675; Glyco_trans_29.
CC Pfam; PF00777; Glyco_transf_29; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
KW

```

```

FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 9 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 30 336 LUMENAL, CATALYTIC (POTENTIAL).
FT DOMAIN 38 49 POLY-GLN.
FT DISULFID 96 245 BY SIMILARITY.
FT CARBOHYD 137 161 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 336 AA; 38443 MW; F5FDD43D45CALICE CRC64;

Alignment Scores:
Pred. No.: 1,07e-22 Length: 336
Score: 328.50 Matches: 65
Percent Similarity: 59.30% Conservative: 37
Best Local Similarity: 37.79% Mismatches: 61
Query Match: 13.79% Indels: 9
DB: 1 Gaps: 2

US-09-714-936-218 (1-1294) x SI7E_HUMAN (1-336)
QY 289 CAACCTGTGACAAAGTGGATACCATCTCCATACATACAGGGGGCC----- 336
Dbb 57 GlnProAlaLaGlu-----SerSerThrGlnGlnArgProGlyValProAla 72
QY 337 ---CCTCGAATCACTACTGATACATATAATGTGAAGACACAGAGCCTTTCACCTGGAC 393
Dbb 73 GlyProArgProLeuAspGlyTyrLeuGlyValAlaAlaAspHisLysProLeuLysMetHis 92
QY 394 TGTGACCTTTGTCATAGTCAAACTCAGTCAGTGCAGATGTGGCCAGAGGTGGGAAT 453
Dbb 93 CysArgAspCysAlaLeuValThrSerSerGlyHisLeuHisSerArgGlnGlySer 112
QY 454 GAGATAGATCATCTCTCTGATTTGGAGATGAACATGCCGCCACCAAAAGGTTAGAA 513
Dbb 113 GlnIleAspGlnThrGlnCysValIleArgMetAsnAspAlaProThrArgGlyTyrGly 132
QY 514 GAAGATGTGGCCGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 573
Dbb 133 ArgAspValGlyAsnArgThrSerLeuArgValIleAlaHisSerSerIleGlnArgile 152
QY 574 CTAAGAAACCCGTATTATTTTCAAGGAGCGAATACTACTATTGTGTATTGGGA 633
Dbb 153 LeuArgAsnArgHisAspLeuLeuAsnValSerGlnGlyThrValPheIlePheTrpGly 172
QY 634 CTTTCCCGCAATAGCAAGATGCAATGGCATCGTTTACAAATGTTTGAAGAGACA 693
Dbb 173 ProSerSerTyrMetArgAspGlyGlyGlnValTyrAsnAsnLeuHisLeuLeu 192
QY 694 GTGTGATCTATCCGAATGCCAAATATACGTGACACAGAGCGCATGAGTTACTGT 753
Dbb 193 SerGlnValLeuProArgLeuLysAlaPheMetIleThrArgHisLysMetLeuGlnPhe 212
QY 754 GATGGAGTATTTAAGAGAGAACTGGGAGGACAGG 789
Dbb 213 AspGluLeuPheLysGlnGlnThrGlyLysAspArg 224

RESULT 6
SI7E_MOUSE
ID SI7E_MOUSE STANDARD; PRT; 336 AA.
AC Q90YJ1; Q90K6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase V (GDI alpha
DE synthase) (GalNAc alpha-2,6-sialyltransferase V) (ST6GalNac V)
DE (Sialyltransferase 7E).
GN SIAT7E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

```

```

RC STRAIN-C3H/HeN, C57BL/6N, and C57BL/6J; TISSUE=Brain;
RX MEDLINE=20069453; PubMed=10601645;
RA Ikehara Y., Shimizu N., Kono M., Nishihara S., Nakanishi H.,
RA Kitamura T., Narimatsu H., Tsuji S., Tatematsu M.;
RT "A novel glycosyltransferase with a polyglutamine repeat; a new
RT candidate for GDIalpha synthase (ST6GalNac V).";
RL FEBS Lett. 463:92-96(1999).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=99452943; PubMed=10521438;
RA Okajima T., Fukumoto S., Ito H., Kiso M., Hirabayashi Y., Urano T.,
RA Furukawa K.;
RT "Molecular cloning of brain specific GDIalpha synthase (ST6GalNac V)
RT containing CAG/glutamine repeats.";
RL J. Biol. Chem. 274:30557-30562(1999).
CC -!- FUNCTION: Involved in the biosynthesis of ganglioside GD1a from
CC Gmb. It exhibits higher activity with glycolipids than with
CC glycoproteins.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
CC -!- TISSUE SPECIFICITY: HIGH EXPRESSION IN FOREBRAIN AND TO A LESSER
CC EXTENT IN CEREBELLUM. NO EXPRESSION IN SALIVARY GLAND, INTESTINE,
CC LIVER, KIDNEY, HEART, LUNG, THYMUS AND SPLEEN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB028840; BAA89292.1; -
CC EMBL; AB030836; BAA85747.1; -
CC MGd; MG1:1349471; Slat7e.
CC InterPro; IPR001675; Glyco.trans.29.
CC Pfam; PF00777; Glyco.transf.29; 1.
CC Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 9 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
CC DOMAIN 30 336 LUMENAL, CATALYTIC (POTENTIAL).
CC DOMAIN 38 49 POLY-GLN.
CC DISULFID 96 245 BY SIMILARITY.
CC CARBOHYD 137 161 N-LINKED (GLCNAC... ) (POTENTIAL).
CC CARBOHYD 161 161 N-LINKED (GLCNAC... ) (POTENTIAL).
CC VARIANT 48 49 MISSING (IN STRAIN C57BL/6J).
CC VARIANT 49 49 MISSING (IN STRAINS C57BL/6N AND
CC C57BL/6J).
SQ SEQUENCE 336 AA; 38430 MW; 128516E3815985E6 CRC64;

Alignment Scores:
Pred. No.: 7,49e-22 Length: 336
Score: 319.50 Matches: 68
Percent Similarity: 58.52% Conservative: 35
Best Local Similarity: 38.64% Mismatches: 64
Query Match: 13.41% Indels: 9
DB: 1 Gaps: 2

US-09-714-936-218 (1-1294) x SI7E_MOUSE (1-336)
QY 278 ACTGCTTTGGACAAACCTGGTACAAAGTGATACCATCTTCCT-----ACACAT 325
Dbb 54 ThrGlySerThrGlnLeuValGlnSer-----SerProGlnProArgGlyThrAla 70
QY 336 ACAGCGCGCCCTTGCAGCTCACTGATGATGATGATGATGATGATGATGATGATGATG 385
Dbb 71 ProAlaGlyProArgGlnLeuGlu--GlyTyrLeuGlyValAlaAspHisLysProLeuL 90
QY 386 RACTGGACTGTGACCTTTGTGCCATAGTGTCAAACTCAGTTCAGATGTTGGCCAGAAGG 445

```

```

Db      90 ysmethiscysylsaspysalaleuvalnrserserclgylhisleuLeuArgSerGing 110
QY      446 TGGGAATGAGATAGATGATCCCTCCGATTTGGAGAAATGAACAATGCCCCACCAAAAG 505
Db      110 lnglyprohis1leaspGlnThrGlycysValleargMetasnaSpAlapProThArgG 130
QY      506 GTTATGAAGAAGATGCGGCGCATGACCATGATTCGAGTTCGATGTCCTCCATACAGCGTTC 565
Db      130 lyTyrglyLeuaspValglyasnArgThrSerLeuArgValleAlaHisSerSerileg 150
QY      566 CTCTTTTGTATAAAGACCTGATTATTTTTCAGGAAGCGAATACATGATTATTGTGTTA 625
Db      150 lngargileuargasnarghisaspLeuLeuasnValSerGlnGlyThrValpHeilep 170
QY      626 TTTGGGACCTTCCGCAATATGAGGAAGATGGCAATGGCATGCTGTTACAACATGTTGA 685
Db      170 heTrpGlyProSerSerTyrrMetArgArgaspGlyLysGlyGlnAlaTyrrAsnAsnLeuG 190
QY      686 AAAAGACAGTTGGTATCTATCCGATGCCAATATACGTGACACAGAGACGCGATGA 745
Db      190 lnLeuLeuSerGlnValLeuProargLeuLysAlaPheMetIleThrArgHisArgMetL 210
QY      746 GTTACTGTGATGAGTGTTTTAAAGAAAGAACTGGGAAGGACAGG 789
Db      210 euGlnPheaspGluLeuPheLysGlnGluThrGlyLysaspArg 224

```

RESULT 7

```

SIAL_CHICK
ID SIAL_CHICK STANDARD; PRT; 413 AA.
AC Q92182;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminide-beta-galactosamide-alpha-2,6-sialyltransferase
DE (EC 2.4.99.1) (Beta-galactoside alpha-2,6-sialyltransferase)
DE (Alpha 2,6-ST) (Sialyltransferase 1) (ST6Gal I).
GN SIALT1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94139712; PubMed=8307003;
RA Kurosawa N., Kawasaki M., Hamamoto T., Nakaoka T., Lee Y.-C.,
RA Arita M., Tsuji S.;
RT "Molecular cloning and expression of chick embryo Gal beta 1,4GlcNAc
RL Eur. J. Biochem. 219:375-381(1994).
CC -!- FUNCTION: TRANSFERS SIALIC ACID FROM THE DONOR OF SUBSTRATE CMP-
CC SIALIC ACID TO GALACTOSE CONTAINING ACCEPTOR SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminide + beta-D-galactosyl-
CC 1,4-acetyl-beta-D-glucosamine = CMP + alpha-N-acetylneuraminyl-
CC 2,6-beta-D-galactosyl-1,4-N-acetyl-beta-D-glucosamine.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI, SOLUBLE FORM IN BODY FLUIDS.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTES, BRAIN AND LIVER
CC AND TO A LESSER EXTENT IN LUNG AND HEART.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

```

CC      EMBL; X75558; CAAS23235.1; -.
DR      PIR; S41114; S41114.
DR      InterPro; IPR001675; Glyco.trans.29.
DR      Pfam; PF00777; Glyco.transf.29; 1.
KW      Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW      Signal-anchor; Golgi stack.
FT      DOMAIN 1 9
FT      TRANSMEM 10 26
FT      SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT      (POTENTIAL).
FT      LUMENAL, CATALYTIC (POTENTIAL).
FT      BY SIMILARITY.
FT      DISULFID 27 413
FT      CARBOHYD 129 342
FT      CARBOHYD 129 329
FT      CARBOHYD 155 155
FT      CARBOHYD 167 167
FT      CARBOHYD 181 181
FT      CARBOHYD 181 181
SQ      SEQUENCE 413 AA; 47392 MW; 19B8CCD361ED137D CRC64;

```

Alignment Scores:

```

Pred. No.: 3,89e-07 Length: 413
Score: 163.00 Matches: 47
Percent Similarity: 40.68% Conservative: 25
Best Local Similarity: 26.55% Mismatches: 47
Query Match: 6.84% Indels: 59
DB: 1 Gaps: 4

```

US-09-714-936-218 (1-1294) x SIAL_CHICK (1-413)

```

QY      343 ACTCACTATGGATACATAAATGTGAAGACACAAAGCCGTTTCAACTGGAGTGTGACCTT 402
Db      131 ThrTyraSngGlyLysMetAsnAlaLysLeuSerProGluGlnLeuLeuCysArgLeu 150
QY      402 -----
Db      151 ArgAspArgValAsnValThrMetIleArgGlySerAspGlyProPheAsnSerSerGlu 170
QY      403 -----TGT 405
Db      171 TrpGlnHisTyrrLeuProAspLysSerLeuAsnGlnThrValGlyArgLeuGlyArgCys 190
QY      406 GCCATATGTCACAACTCAAGTCAGGTGCAGATGGTTGGCCAGAGGTGGGAATAGATAGATCGA 465
Db      191 AlaValValSerSerAlaGlySerLeuLysSerSerHisLeuGlyProGluIleAspSer 210
QY      466 TCCTCTGCGATTTGGAGATGAACAATGCCCCACCAAGAGTTATGAGAGATGTCGCG 525
Db      211 HisAspAlaValLeuArgPheAsnGlyAlaProValLysGlyPheGlnGluAspValGly 230
QY      526 CGCATGACCATGATTCGAGTTGTCTCCATACCCAGCGTTCCTCTTTTGTCTAAAAAACCCCT 585
Db      231 GlnLysThrThrIleArgLeuValAsnSer-----GlnLeuValThrValGluGlu 247
QY      586 GATTATTTTTCAGGAAGCG-----AATACATCTATTTGTGTATTGTTGGGACCTTT- 638
Db      248 GlnGlnPheLeuLysAspAlaLeuTyrrAsnThrGlyIleLeuIleValTrpAspProAla 267
QY      639 CCCCAATATGAGAAAGATGCAATGG-----
Db      268 ProTyrrHisAlaGlnIleHisGlnIleTyrrArgLysProAspTyrrLysPhePheGluAla 287
QY      666 -----CATCGTTTACACATGTTGAAAGACAGAGTGTGATCTATCC 707
Db      288 TyrrLysSerTyrrArgIleArgHisProGluGlnProPheTyrrIleLeuAsn 304

```

RESULT 8

```

SIAL_MOUSE
ID SIAL_MOUSE STANDARD; PRT; 403 AA.
AC Q64685;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminide-beta-galactosamide-alpha-2,6-sialyltransferase
DE (EC 2.4.99.1) (Beta-galactoside alpha-2,6-sialyltransferase)

```

DE (Alpha 2,6-ST) (Sialyltransferase 1) (ST6Gal I).
GN SIAT1.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Liver;
RX MEDLINE=94363344; PubMed=8081843;
RA Hamamoto T., Kawasaki M., Kurosawa N., Nakaoka T., Lee Y.-C.,
RA Tsuji S.;
RT "Two step single primer mediated polymerase chain reaction.
RT Application to cloning of putative mouse, beta-galactoside alpha 2,6-
RL sialyltransferase cDNA.";
RL Bioorg. Med. Chem. 1:141-145(1993).
CC -!- FUNCTION: TRANSFERS SIALIC ACID FROM THE DONOR OF SUBSTRATE CMP-
CC SIALIC ACID TO GALACTOSE CONTAINING ACCEPTOR SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminate + beta-D-galactosyl-
CC 1,4-acetyl-beta-D-glucosamine = CMP + alpha-N-acetylneuraminy-
CC 2,6-beta-D-galactosyl-1,4-N-acetyl-beta-D-glucosamine.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI, SOLUBLE FORM IN BODY FLUIDS.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; D16106; BAA03680.1; -;
DR MGD; MGI:108470; Siat1.
DR InterPro; IPR001675; Glyco.trans.29.
DR Pfam; PF00777; Glyco.transf.29; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT
FT DOMAIN 27 403 LUMENAL, CATALYTIC (POTENTIAL).
FT DISULFID 181 332 BY SIMILARITY.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 403 AA; 46407 MW; ECARD0603A5A947B CRC64;

Alignment Scores:
Pred. No.: 1.37e-05 Length: 403
Score: 146.50 Matches: 32
Percent Similarity: 42.48% Conservative: 16
Best Local Similarity: 28.32% Mismatches: 36
Query Match: 6.15% Indels: 29
DB: 1 Gaps: 1

US-09-714-936-218 (1-1294) x SIAL_MOUSE (1-403)

QY 532 ACCATGATTCGAGTGTGTCC----- 552
 ||| |||||:::||||:
 Db 224 ThrThrIleArgLeuValAsnSerGlnLeuValThrGluLysArgPheLeuLysAsp 243
 QY 553 -----CATACCAGCGT 564
 |||
 Db 244 SerLeuTyrThrGluGlyIleLeuIleLeuTrpAspProSerValTyrHisAlaAspIle 263
 QY 565 CCTCTTTTGTAAACACCTGATTATTTTTCACAGGA 603
 ||| :: ||||| ||| |||
 Db 264 ProGlnTrpTyrGlnLysProAspTyrAsnPhePheGlu 276
 RESULT 9
 SIAL_HUMAN STANDARD; PRT; 406 AA.
 AC P15907;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,6-sialyltransferase
 DE (EC 2.4.99.1) (Beta-galactoside alpha-2,6-sialyltransferase)
 DE DE (Alpha 2,6-ST) (Sialyltransferase 1) (ST6Gal I) (B-cell antigen CD75).
 GN SIAT1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=90175005; PubMed=2408023;
 RA Grundmann U.G., Nerlich C., Rein T., Zettlmeissl G.;
 RT "Complete cDNA sequence encoding human beta-galactoside alpha-2,6-
 RT sialyltransferase.";
 RT Nucleic Acids Res. 18:667-667(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90324884; PubMed=2373995;
 RA Stamenkovic I., Asheim H.C., Deggerdal A., Blomhoff H.K.,
 RA Smeland E.B., Funderud S.;
 RT "The B cell antigen CD75 is a cell surface sialyltransferase.";
 RT J. Exp. Med. 172:641-643(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=92112983; PubMed=1730763;
 RA Bast B.J.E.G., Zhou L.J., Freeman G.J., Colley K.J., Ernst T.J.,
 RA Munro J.M., Tedder T.F.;
 RT "The HB-6, CD75, and CD76 differentiation antigens are unique cell-
 RT surface carbohydrate determinants generated by the beta-galactoside
 RT alpha 2,6-sialyltransferase.";
 RT J. Cell Biol. 116:423-435(1992).
 RN [4]
 RP SEQUENCE OF 74-406 FROM N.A.
 RX MEDLINE=90026381; PubMed=2803295;
 RA Lance P., Lau K.M., Lau J.T.Y.;
 RT "Isolation and characterization of a partial cDNA for a human
 RT sialyltransferase.";
 RT Biochem. Biophys. Res. Commun. 164:225-232(1989).
 CC -!- FUNCTION: TRANSFERS SIALIC ACID FROM THE DONOR OF SUBSTRATE CMP-
 CC SIALIC ACID TO GALACTOSE CONTAINING ACCEPTOR SUBSTRATES.
 CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminate + beta-D-galactosyl-
 CC 1,4-acetyl-beta-D-glucosamine = CMP + alpha-N-acetylneuraminy-
 CC 2,6-beta-D-galactosyl-1,4-N-acetyl-beta-D-glucosamine.
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI, SOLUBLE FORM IN BODY FLUIDS.
 CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.
 CC -!- PTM: THE HB-6, CD75, AND CD76 DIFFERENTIATION ANTIGENS ARE CELL-
 CC SURFACE CARBOHYDRATE DETERMINANTS GENERATED BY THIS ENZYME.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
 CC -!- DATABASE: NAME=PROW; NOTE=PROW 2:23-31(2001);


```

FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 232 MHNPKKFSILFLVFAVICVWKSGSDYDALFLOAKE
FT KVKPKSQFQVWDKSTYKLNPRLLKRWYLNANKYKVS
FT KGPQGVKESVENLRCHLDHVNYSMTIARDFFPNTIEWEG
FT YLRFNPKFKVQWQKAVSSAGSLNSOLGREIDNHDAV
FT LRFNGAPTDNFOODVGSKTIRLMNSQ -> MRYLLFWYGL
FT PHYSQCVCWTPASGIFENFLLSLLLVLGK (in
FT isoform RKA).
FT /FTid=VSP_001782.
FT MHNPKKFSILFLVFAVICVWKSGSDYDALFLOAKE
FT FOMPKSQEKVAGSASQVFNFSQDKPEDIPLSHRVTA
FT KVKPKSQFQVWDKSTYKLNPRLLKRWYLNANKYKVS
FT KGPQGVKESVENLRCHLDHVNYSMTIARDFFPNTIEWEG
FT YLRFNPKFKVQWQKAVSSAGSLNSOLGREIDNHDAV
FT LRFNGAPTDNFOODVGSKTIRLMNSQ -> MRYLLFWYGL
FT PHS (in isoform RKA).
FT /FTid=VSP_001783.
SQ SEQUENCE 403 AA; 46732 MW; F40D7EA6B67EA9A CRC64;

Alignment Scores:
Pred. No.: 7.7e-05
Score: 138.50
Percent Similarity: 45.19%
Best Local Similarity: 23
Query Match: 5.81%
Indels: 27
Gaps: 6

US-09-714-936-218 (1-1294) x SIAL_RAT (1-403)
QY 238 GTGGCTGTGAATGAAGTGAATTCCTCACTTAAGTCTTGGACACCTGCT 297
Db 147 ValSerMetIleClnAlaThrAspPheProPhe-----AsnThr 159
QY 298 ACAAGTGGATACCATTCCTCATACATACAGCGGCCCTTCGAAGTCACTCATGATAC 357
Db 160 ThrGluTrp-----GluGlyTyLeuProLysGluAsnPheArgThrLysValGly--- 176
QY 358 ATAAAGTGAACACAGAGCCCTTCGACTGAGTGTGACCTTTGGCCAGTGTCA 417
Db 177 -----ProTirpGln-----ArgCysAlaValValSer 185
QY 418 AACTCAGTCAGATGTTGGCCAGAGAGTGGAAAGAGATAGATCGATCCTCGCAT 477
Db 186 SerAlaGlySerLeuLysAsnSerGlnLeuGlyArgGluLysAsnHisAspAlaVal 205
QY 478 TGGAGATGAACAATGCCCCACCAAGGTTATGAAGAGATGTGGCCGCGCATGACCAT 537
Db 206 LeuArgPheAsnGlyAlaProThrAspAsnPheGlnAspValGlySerLysThrThr 225
QY 538 ATTGCGATGTCGTCATACACGAGTTCCTTTCTTAAACAAACCCCTGATTATTTC 597
Db 226 IleArgLeuMetAsnSerGln-----LeuValThrThrGluLysArgPheLeu 241
QY 598 AAGGAAGCAATCACTACT-----ATTGTGTATTGGGACCT 636
Db 242 LysAspSerLeuTyThrGluGlyLeuLeuIleValTrpAspPro 256

RESULT 12
SIA6_MOUSE
ID SIA6_MOUSE STANDARD; PRN; 374 AA.
AC P97325;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminat-beta-1,4-galactoside alpha-2,3-
DE sialyltransferase (EC 2.4.99.6) (N-acetylglucosaminide alpha-2,3-
DE sialyltransferase) (Gal beta-1,3(4) GlcNAc alpha-2,3
DE sialyltransferase) (ST3N) (Sialyltransferase 6).
GN SIA6 OR SIAF3.
OS Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain, and Liver;
MEDLINE=97328289; PubMed=9184827;
RA Kono M., Ohshima Y., Lee Y.-C., Hamamoto T., Kojima N., Tsuji S.;
RT "Mouse beta-galactoside alpha2,3-sialyltransferases: comparison of
in vitro substrate specificities and tissue specific expression.";
RL Glycobiology 7:469-479(1997).
CC -!- FUNCTION: CATALYZES THE FORMATION OF THE NEURAC-ALPHA-2,3-GAL-BETA-
CC 1,4-GLCAC, NEURAC-ALPHA-2,3-GAL-BETA-1,3-GLCNAC OR NEURAC-ALPHA-
CC 2,3-GAL-BETA-1,3-GALNAC SEQUENCES FOUND IN TERMINAL CARBOHYDRATE
CC GROUPS OF GLYCOPROTEINS AND GLYCOLIPIDS. THE HIGHEST ACTIVITY IS
CC TOWARD GAL-BETA-1,3-GLCNAC AND THE LOWEST TOWARD GAL-BETA-1,3-
CC GALNAC.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminat + beta-D-galactosyl-
CC acetylneuraminyl-2,3-beta-D-galactosyl-1,4-N-acetyl-D-
CC glucosaminyl-glycoprotein.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: FOUND IN ALL TISSUES TESTED. HIGH EXPRESSION
CC FOUND IN BRAIN, LIVER, KIDNEY, COLON, HEART AND SPLEEN.
CC -!- DEVELOPMENTAL STAGE: DEVELOPMENTAL REGULATION ONLY OCCURS IN
CC LIVER, HEART, KIDNEY AND SPLEEN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X84234; CAA59013.1; .
CC MGD; MGI:1316659; Siat6.
CC InterPro; IPR001675; Glyco.trans.29.
CC Pfam; PF00777; Glyco.transf.29; 1.
CC Transference; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 9 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC
CC DOMAIN 29 374 LUMENAL, CATALYTIC (POTENTIAL).
CC DISULFID 159 313 BY SIMILARITY.
CC CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 374 AA; 42131 MW; C6F2BA5189C29532 CRC64;

Alignment Scores:
Pred. No.: 9.35e-05
Score: 137.50
Percent Similarity: 39.78%
Best Local Similarity: 21.55%
Query Match: 5.77%
Indels: 45
Gaps: 4

US-09-714-936-218 (1-1294) x SIA6_MOUSE (1-374)
QY 283 TTTGGACAACCTGGTACAAAGTGGATACCATCTCTCATACATACAGCGGCCCTTCGA 342
Db 101 PheSerLysProAlaProMetPheLeuAspSerPheArgLysTrpAlaArgIleArg 120
QY 343 ACTCACTATGTATACATAAATGTGAAGACACACAGACCTTTG----- 384
Db 121 GluPheValProPheGlyLeuLysGlyGlnAspAsnLeuIleLysAlaIleLeuSer 140
QY 385 -----CACTGGACTGTGACCTTTGTCGC 408
Db 141 ValThrLysGluTyArgLeuThrProAlaLeuAspSerLeuHisCysArgArgCysIle 160

```



```

Db      256 Lys 256
RESULT 14
SI4A_CHICK
ID SI4A_CHICK STANDARD; PRT; 342 AA.
AC Q11200;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminatase-beta-galactosaminide-alpha-2,3-sialyltransferase
DE (EC 2.4.99.4) (Beta-galactoside alpha-2,3-sialyltransferase) (Alpha
DE 2,3-ST) (Gal-NAC6S) (Gal-beta-1,3-GalNAc-alpha-2,3-sialyltransferase)
DE (ST3GALIA) (ST30) (ST3GALIA.1) (SIAT4-A) (ST3Gal I).
GN SIAT4A.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic brain;
RA MEDLINE=95284088; PubMed=7766661;
RA Kurosawa N., Hamamoto T., Inoue M., Tsuji S.;
RT "Molecular cloning and expression of chick Gal beta 1,3GalNAc alpha
RT 2,3-sialyltransferase."
RL Biochim. Biophys. Acta 1244:216-222(1995).
CC -!- FUNCTION: IT MAY BE RESPONSIBLE FOR THE SYNTHESIS OF THE SEQUENCE
CC NEUAC-ALPHA-2,3-GAL-BETA-1,3-GALNAc- FOUND ON SUGAR CHAINS
CC O-LINKED TO THR OR SER AND ALSO AS A TERMINAL SEQUENCE ON CERTAIN
CC GANGLIOSIDES. SIAT4-A AND SIAT4-B SIALYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminatase + beta-D-galactosyl-
CC 1,3-N-acetyl-alpha-D-galactosaminyl-R = CMP + alpha-N-
CC acetylneuraminyl-2,3-beta-D-galactosyl-1,3-N-acetyl-alpha-D-
CC galactosaminyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI, SOLUBLE FORM IN BODY FLUIDS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN EARLY EMBRYONIC STAGES.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X80503; CAA56666.1; -.
CC PIR; S55675; S55675.
CC InterPro; IPR001675; Glyco.trans.29.
CC Pfam; PF00777; Glyco.transf.29.1
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 11 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC DOMAIN 29 342 LUMENAL, CATALYTIC (POTENTIAL).
CC DISULFID 144 283 BY SIMILARITY.
CC CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 342 AA; 39540 MW; 59B657652F4FE949 CRC64;

```

```

Alignment Scores:
Pred. No.: 0.000126 Length: 342
Score: 136.00 Matches: 42
Percent Similarity: 44.36% Conservative: 17
Best Local Similarity: 31.58% Mismatches: 54
Query Match: 5.71% Indels: 20
DB: 1 Gaps: 4

US-09-714-936-218 (1-1294) x SI4A_CHICK (1-342)
QY 373 CAAGAGCCCTTGCAC-----CTGACTGTGACCTTTGTGCCATAGTCTCAAC 420
Db 130 ArgAspProLeuGlnGluArgGlyThrPheSerCysArgGlyCysAlaValValGlyAsn 149
QY 421 TCAGGTCAGATGGTTGGCCAGAGGTGGAAATACATAGATCGATCCTCTCGATTGG 480
Db 150 SerGlyAsnLeuArgGlnSerGlnTyrGlyGlnAspLeuSerHisAspPheValLeu 169
QY 481 AGAATGAACAATGCCCCACCAAGGTTATGAAGAAGATGCGCGCCGATGACCATGATT 540
Db 170 ArgMetAsnArgAlaProThrIleGlyTyrGluSerAspValGlySerLysThrHis 189
QY 541 CGAGTTGTGTCACACACAGCGTTCCTCTTTGCTAAAAAACCTGATATTATTTTCAAG 600
Db 190 HisPheValTyrProGluSer-----TyrLys 198
QY 601 GAACGGATACTACTATTGTGTTATTGGGACCTTCCGCAATATGAGGAAGATGGC 660
Db 199 GluLeuAlaGluAsnValSerMetIleValIleProPheLysThrLeu-----AspLeu 216
QY 661 AATGGCATCTTTTACACATGTTGAAAAAGACAGTGGT-----ATCTATCCGAAT 711
Db 217 ArgTrpIleValThrAlaLeuThrThrGlyThrIleAsnPheThrTyrValProValPro 236
QY 712 GCCCAATATACGTGACACAGAGAGAGCGCATGAGTTAC 750
Db 237 ArgLysIleLysValArgLysGluLysValLeuIleTyr 249

RESULT 15
SI4B_MOUSE
ID SI4B_MOUSE STANDARD; PRT; 350 AA.
AC Q11204;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminatase-beta-galactosaminide-alpha-2,3-sialyltransferase
DE (EC 2.4.99.-) (Beta-galactoside alpha-2,3-sialyltransferase) (Alpha
DE 2,3-ST) (Gal-NAC6S) (Gal-beta-1,3-GalNAc-alpha-2,3-sialyltransferase)
DE (ST3GALIA.2) (SIAT4-B) (ST3Gal II).
GN SIAT4B OR SIAT5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Schiuognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=94193584; PubMed=8144500;
RA Lee Y.-C., Kojima N., Wada E., Kurosawa N., Nakaoka T., Hamamoto T.,
RA Tsuji S.;
RT "Cloning and expression of cDNA for a new type of Gal beta 1,3GalNAc
RT alpha 2,3-sialyltransferase."
RL J. Biol. Chem. 269:10028-10033(1994).
CC -!- FUNCTION: IT MAY BE RESPONSIBLE FOR THE SYNTHESIS OF THE SEQUENCE
CC NEUAC-ALPHA-2,3-GAL-BETA-1,3-GALNAc- FOUND IN TERMINAL CARBOHYDRATE
CC GROUPS OF CERTAIN GLYCOPROTEINS, OLIGOSACCHARIDES AND GLYCOLIPIDS.
CC SIAT4-A AND SIAT4-B SIALYLATE THE SAME ACCEPTOR SUBSTRATES BUT
CC EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: CMP-N-acetylneuraminatase + beta-D-galactosyl-
CC 1,3-N-acetyl-alpha-D-galactosaminyl-R = CMP + alpha-N-
CC acetylneuraminyl-2,3-beta-D-galactosyl-1,3-N-acetyl-alpha-D-
CC galactosaminyl-R.
CC -!- PATHWAY: Glycosylation.

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:32:19 ; Search time 21 Seconds

(without alignments)
5214.310 Million cell updates/sec

Title: US-09-714-936-218

Perfect score: 2383

Sequence: 1 ccggaatttcgggtgcagc.....ttctctctctttttttg 1294

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=xlh

-O=/cgn2_1/USPTO.spool/US09714936/runat_10092003_151500_5727/app_query.fasta_1.1479

-DB-Issued Patents AA -QRM-fastan -SUFFIX=n2p.rai -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09714936.ecgn_1.1.18.@runat_10092003_151500_5727 -NCPU=6 -ICPU=3

-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/6C_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382	16.0	302	4	US-09-599-360B-75
2	328.5	13.8	336	3	US-09-334-601-13
3	146.5	6.1	376	2	US-08-666-367B-8
4	146.5	6.1	376	3	US-09-143-438-8
5	142	6.0	406	1	US-08-446-777-4
6	142	6.0	767	1	US-08-446-777-6
7	142	6.0	767	1	US-08-446-777-8
8	141.5	5.9	90	3	US-09-334-601-20
9	141.5	5.9	90	3	US-09-334-601-21
10	140.5	5.9	329	3	US-09-334-601-10
11	138.5	5.8	90	3	US-09-334-601-22
12	138.5	5.8	340	2	US-08-102-385G-18

13	137.5	5.8	375	2	US-08-446-875-10	Sequence 10, Appl
14	137.5	5.8	375	2	US-08-102-385G-10	Sequence 10, Appl
15	135.5	5.7	340	2	US-08-446-875-16	Sequence 16, Appl
16	135.5	5.7	374	2	US-08-446-875-4	Sequence 4, Appl
17	135.5	5.7	374	2	US-08-102-385G-4	Sequence 4, Appl
18	134.5	5.6	384	3	US-09-334-601-8	Sequence 8, Appl
19	133.5	5.6	90	3	US-09-334-601-26	Sequence 26, Appl
20	129.5	5.4	333	1	US-07-991-587A-2	Sequence 2, Appl
21	129.5	5.4	333	1	US-08-309-985-2	Sequence 2, Appl
22	125.5	5.3	343	2	US-08-446-875-2	Sequence 2, Appl
23	125.5	5.3	343	2	US-08-102-385G-2	Sequence 2, Appl
24	124.5	5.2	92	3	US-09-334-601-29	Sequence 29, Appl
25	124	5.2	339	1	US-08-626-994A-3	Sequence 3, Appl
26	124	5.2	339	3	US-08-957-742-3	Sequence 3, Appl
27	124	5.2	364	1	US-08-626-994A-1	Sequence 1, Appl
28	124	5.2	364	3	US-08-957-742-1	Sequence 1, Appl
29	123	5.2	90	3	US-09-334-601-24	Sequence 24, Appl
30	123	5.2	92	3	US-09-334-601-18	Sequence 18, Appl
31	123	5.2	92	3	US-09-334-601-19	Sequence 19, Appl
32	122	5.1	92	3	US-09-334-601-23	Sequence 23, Appl
33	122	5.1	92	3	US-09-334-601-17	Sequence 17, Appl
34	121	5.1	55	2	US-08-102-385G-14	Sequence 14, Appl
35	120	5.0	90	3	US-09-334-601-25	Sequence 25, Appl
36	119	5.0	77	2	US-08-102-385G-30	Sequence 30, Appl
37	119	5.0	359	4	US-09-425-488-2	Sequence 2, Appl
38	119	5.0	414	3	US-09-334-601-4	Sequence 4, Appl
39	118.5	5.0	329	1	US-07-991-587A-7	Sequence 7, Appl
40	118.5	5.0	329	1	US-08-309-985-7	Sequence 7, Appl
41	118.5	5.0	332	2	US-08-446-875-12	Sequence 12, Appl
42	118.5	5.0	332	2	US-08-102-385G-12	Sequence 12, Appl
43	118	5.0	55	2	US-08-102-385G-16	Sequence 16, Appl
44	117.5	4.9	92	3	US-09-334-601-28	Sequence 28, Appl
45	116.5	4.9	290	3	US-09-068-655-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-09-599-360B-75

; Sequence 75, Application US/09599360B

; Patent No. 6548633

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Bougueleret, L.

; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides

; FILE REFERENCE: GENSET.050CP3

; CURRENT APPLICATION NUMBER: US/09/599,360B

; CURRENT FILING DATE: 2000-06-21

; PRIOR APPLICATION NUMBER: 60/113,686

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: 60/141,032

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 09/469,099

; PRIOR FILING DATE: 1999-12-21

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: Patent.pm

; SEQ ID NO 75

; LENGTH: 302

; TYPE: PRT

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -18..-1

US-09-599-360B-75

Alignment Scores:

Pred. No.:	9,17e-34	Length:	302
Score:	382.00	Matches:	86
Percent Similarity:	56.57%	Conservative:	37
Best Local Similarity:	40.95%	Mismatches:	59
Query Match:	16.03%	Indels:	28
DB:	4	Gaps:	6

US-09-714-936-218 (1-1294) x US-09-334-601-13 (1-302)

```

QY 230 TGCTGG-----TTGTGCGCTCTGTAAATGAGTGAATTTCCCATGCTACTAA 277
Db 26 CysirpalaglyLeuProleucysleuAala-----35
QY 278 ACTGCTTTGGACAACCTGCTACAAAAGTGCATACCATTTCTCTACATACACA-----328
Db 36 ThrCysLeuasp-----HisHisPheProThrGlySerArgProThr 49
QY 329 ---GGCGGCCCTCGAAGTCTACTATGATGATACATAAATGGAAGACACAGAGCTTTGC 385
Db 50 ValProglyProleuHisPheSer--GlyTyrSerValProaspGlylysProLeuV 69
QY 386 AACTGGAC--TGTGACCTTTGTGCCATGAGTCAACACTCAGTCAGATGTTGGCCAGA 442
Db 69 alaRgGluProCysArgSerCysAlaValSerSerGlyGlnMetLeuGlySerG 89
QY 443 AGGTGGGAATGAGATAGATCGATCTCTGTCATTTGGAGATGAACATGCCCCACCA 502
Db 89 lyLeuGlyAlaGluileaspSerAlaGluCysValPheArgMetasGlnlalaProThrV 109
QY 503 AAGTTATGAAAGATGTCGGCCCGCATGACATGATGAGTGTGTCCTCCATCCAGCG 562
Db 109 alglyPheGluAlaaspValGlyGlnArgSerThrLeuArgValSerHisThrSerV 129
QY 563 TTCTCTTTTGTCTAAAAACCTGATTTATTTTCAAGGAAGCGAATCTACTATTGTCG 622
Db 129 alProleuLeuArgGlnSerHisTyrPheGlnlysAlaArgaspThrLeuTyrM 149
QY 623 TATTTGGGACCTTTCCCAATATAGGAAGATGCAATGSCATCGTTTACACATGT 682
Db 149 etValTTPGlyGlnGlyArgHisMetaspArgValLeuGlyGlyArgThrYrargThrL 169
QY 683 TGAARAACACAGTGTGTATCTATCCGAATGCCCAATATACGTGACCCACAGAGCGCA 742
Db 169 euLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrPheThrGluArgMetM 189
QY 743 TGAGTTACTGTGAGTGTGTTTAAAGAAAGAACTGGGAGGACAGGGGCGCATCAAGGC 802
Db 189 etAlaTyrCysaspGlnlePheGlnaspGluThrGlylysasnArgGlnInserGlys 209
QY 803 GACTGCTGATTTACAGACACTTTT 828
Db 209 erPheLeu---SerThrGlyTTPphe 216

```

RESULT 2

```

US-09-334-601-13
; Sequence 13, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Kapitonov, Dmitri
; APPLICANT: Yu, Robert
; TITLE OF INVENTION: NOVEL SIALLYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-334-601-13

```

Alignment Scores:
 Pred. No.: 9,02e-28 Length: 336
 Score: 328.50 Matches: 65
 Percent Similarity: 59.30% Conservative: 37
 Best Local Similarity: 37.79% Mismatches: 61
 Query Match: 13.79% Indels: 9
 DB: 2 Gaps: 2

US-09-714-936-218 (1-1294) x US-09-334-601-13 (1-336)

```

QY 289 CAACTGCTGACAAAGGGATACCATCTCTACACATACAGAGGGGCC-----336
Db 57 GlnProAlaLaaglu-----SerSerThrGlnGlnArgProglyValProAla 72
QY 337 --CTTGGAACCTCAGTATGATACATAAATGTGAAGACACAAAGAGCTTGTCACTGGAC 393
Db 73 GlyProArgProleuaspGlyTyrLeuGlyValAlaaspHisLysProleuLysMetHis 92
QY 394 TGTGACCTTTTGCCATGAGTGTCAAACTCAGTCAGATGTTGGCCAGAGGTGGGAAT 453
Db 93 CysArgaspCysAlaLeuValThrSerSerGlyHisLeuLeuHisSerArgGlnGlySer 112
QY 454 GAGATAGATCGATCCCTCCGTCATTTGGAGATGACAAATCCCCCACCACCAAGTTATGAA 513
Db 113 GlnleaspGlnThrGluCysValleargMetasnAspAlaProThrArgGlyTyrGly 132
QY 514 GAAGATGTCGGCCGATGATGATTCGATGTCATACATACAGAGCTTCTCTTTTG 573
Db 133 ArgaspValGlyasnArgThrSerLeuArgValleahisSerSerleleGlnArgIle 152
QY 574 CTAAAAAACCTGATTAATTTTCAAGGAAGCGCAATCTACTATTGTTGTTATTTGGGGA 633
Db 153 LeuArgasnArgHisAspLeuLeuasnValSerGlnGlyThrValPheilePheTrpGly 172
QY 634 CCTTCCCGCATATGAGGAAGATGCGCAATGCGATCGCTTTTACACATGTTGAAAAGACA 693
Db 173 ProSerSerTyrMetArgaspGlyGlyGlnValTyrAsnAsnLeuHisLeuLeu 192
QY 694 GTTGATATCTATCCGAATGCCAATATACGTGACCCACAGAGAGCGCATGAGTTACTGT 753
Db 193 SerGlnValleuProArgTyrLeuLysAlaPheMetileThrArgHisLysMetLeuGlnPhe 212
QY 754 GATGAGTGTTTTAAAGGAAGAACTGGGAAGGACAGG 789
Db 213 AspGluLeuPheLysGlnGluThrGlyLysaspArg 224

```

RESULT 3

```

US-08-666-367B-8
; Sequence 8, Application US/08666367B
; Patent No. 5854042
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
; PRODUCING THE SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,367B
; FILING DATE: August 19, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,567
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850

```


; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 406 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-777-4

Alignment Scores:
Pred. No.: 6.38e-07 Length: 406
Score: 142.00 Matches: 35
Percent Similarity: 49.17% Conservative: 24
Best Local Similarity: 29.17% Mismatches: 41
Query Match: 5.96% Indels: 20
DB: 1 Gaps: 5

US-09-714-936-218 (1-1294) x US-08-446-777-4 (1-406)

```
QY 403 TGTGCAATGAGTCAACTCAGTCAGTGGTGGCCAGAGGTGGGAATGAGATAGAT 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 CysAlaValValSerSerAlaGlySerLeuLysSerGlnLeuGlyArgGluLeasp 203
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 CGATCTCTCTCCTGATGGAGATGACATGCCCCCACCACCAAGTTATGAGAAGATGTC 522
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 AspHisAspAlaValLeuArgPheAsnGlyAlaProThrAlaAsnPheGlnGlnAspVal 223
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 GCGCCGATGACCATGATTCAGTGTGTGCCATACACAGCGTTCCTCTTTTGTCTAAAAAAC 582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 GlyThrLysThrThrIleArgLeuMetAsnSerGln-----LeuValThrThr 239
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 583 CCGTATATATTTTCAAGGAGCG-----AATACTACTATTTGTGTTATTTGGGACCT 636
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 GluLysArgPheLeuLysAspSerLeuTyraAsnGluGlyIleLeuIleValTrpAspPro 259
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 -----TTCCGCAATATGAGGAAGATGCGCAATGGC 666
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 SerValTyHisSerAspIleProLysTrpTyraGlnAsn-----ProAspTyraShn--- 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 ATCGTTTACAACATGTTGAAAAAGACAGTGTGATATCCGAATCCCAATAATACGTG 726
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 277 ---PhePheAsnAsuTyrLysThrTyraArgLysLeuHisProAsnGlnProPheTyrlle 295
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 6
US-08-446-777-6
; Sequence 6, Application US/08446777
; Patent No. 5641668
; GENERAL INFORMATION:
; APPLICANT: Berger, Eric G.
; APPLICANT: Watzele, Manfred
; APPLICANT: Iwanow, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-777-6

Alignment Scores:
Pred. No.: 8.91e-07 Length: 767
Score: 142.00 Matches: 35
Percent Similarity: 49.17% Conservative: 24
Best Local Similarity: 29.17% Mismatches: 41
Query Match: 5.96% Indels: 20
DB: 1 Gaps: 5

US-09-714-936-218 (1-1294) x US-08-446-777-6 (1-767)

```
QY 403 TGTGCAATGAGTCAACTCAGTCAGTGGTGGCCAGAGGTGGGAATGAGATAGAT 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 545 CysAlaValValSerSerAlaGlySerLeuLysSerGlnLeuGlyArgGluLeasp 564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 CGATCTCTCTCCTGATGGAGATGACATGCCCCCACCACCAAGTTATGAGAAGATGTC 522
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 565 AspHisAspAlaValLeuArgPheAsnGlyAlaProThrAlaAsnPheGlnGlnAspVal 584
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 GCGCCGATGACCATGATTCAGTGTGTGCCATACACAGCGTTCCTCTTTTGTCTAAAAAAC 582
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 585 GlyThrLysThrThrIleArgLeuMetAsnSerGln-----LeuValThrThr 600
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 583 CCGTATATATTTTCAAGGAGCG-----AATACTACTATTTGTGTTATTTGGGACCT 636
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 GluLysArgPheLeuLysAspSerLeuTyraAsnGluGlyIleLeuIleValTrpAspPro 620
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 637 -----TTCCGCAATATGAGGAAGATGCGCAATGGC 666
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 621 SerValTyHisSerAspIleProLysTrpTyraGlnAsn-----ProAspTyraShn--- 637
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 ATCGTTTACAACATGTTGAAAAAGACAGTGTGATATCCGAATCCCAATAATACGTG 726
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 638 ---PhePheAsnAsuTyrLysThrTyraArgLysLeuHisProAsnGlnProPheTyrlle 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 7
US-08-446-777-8
; Sequence 8, Application US/08446777
; Patent No. 5641668

;; GENERAL INFORMATION:
;; APPLICANT: Berger, Eric G.
;; APPLICANT: Watzele, Manfred
;; APPLICANT: Ivanow, Svetoslav X.
;; TITLE OF INVENTION: Proteins having glycosyltransferase
;; TITLE OF INVENTION: activity
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CIBA-GEIGY Corporation
;; STREET: 520 White Plains Road
;; CITY: Tarrytown
;; STATE: NY
;; ZIP: 10591-9005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/446,777
;; FILING DATE: May 26, 1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/EP93/03194
;; FILING DATE: 15 Nov 93
;; PRIOR APPLICATION DATA: EPO 92810924.8
;; FILING DATE: 27 Nov 92
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ferraro, Gregory D.
;; REGISTRATION NUMBER: 36,134
;; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 277-3318
;; TELEFAX: (908) 277-4306
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 767 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-446-777-8

Alignment Scores:
Pred. No.: 8,91e-07 Length: 767
Score: 142.00 Matches: 35
Percent Similarity: 49.17% Conservative: 24
Best Local Similarity: 29.17% Mismatches: 41
Query Match: 5.96% Indels: 20
DB: 1 Gaps: 5

US-09-714-936-218 (1-1294) x US-08-446-777-8 (1-767)

QY 403 TGTGCGCATAGTCTCAAACTCAGGTTCAGATGGTTGCCAGAAAGTGGGAATGAGATAGAT 462
Db 545 CysAlaValIserSerAlaGlySerLeuLysSerSerGlnLeuGlyArgGluIleasp 564
QY 463 CGATCTCTCGATTGGAGATGACATGCCCCACCAAGGTTATGAGAAAGATGTC 522
Db 565 AspHisAspAlaValLeuArgPheAsnGlyAlaProThrAlaAsnPheGlnGlnAspVal 584
QY 523 GGCCGCATCATGATCGAGTGTGTCCTCATACCAGCGTCTCTTTGCTTAAAAAC 582
Db 585 GlyThrLysThrIleArgLeuMetAsnSerGln-----LeuValThrThr 600
QY 583 CCTGATTATTTTTCAGGAAGCG-----AATACTACTATTGTGTTATTTGGGACCT 636
Db 601 GlnLysArgPheLeuLysAspSerLeuTyraAsnGlnGlyIleLeuIleValTrpaspPro 620
QY 637 -----TTCCGCAATATGAGGAAGATGCAATGCG 666
Db 621 SerValTyHisSerAspIleProLysTrpTyraGlnAsn-----ProAspTyraAsn--- 637

QY 667 ATCGTTTACACATGTTGAAAAGACAGTGTGTATCTATCTCCGAATGCCAAATATACGTG 726
Db 638 ---PhePheAsnAsnTyrlLysThrTyraGlyLeuHisProAsnGlnProPheTyrlle 656
RESULT 8
US-09-334-601-20
; Sequence 20, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Kapitonov, Dmitri
; APPLICANT: Yu, Robert
; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-334-601-20
Alignment Scores:
Pred. No.: 3,28e-07 Length: 90
Score: 141.50 Matches: 32
Percent Similarity: 52.00% Conservative: 7
Best Local Similarity: 42.67% Mismatches: 29
Query Match: 5.94% Indels: 7
DB: 3 Gaps: 2
US-09-714-936-218 (1-1294) x US-09-334-601-20 (1-90)
QY 394 TGTGACCTTTGTGCGCATAGTGTCAAACTCAGGTTCAGATGGTTGCCAGAAAGTGGGAAT 453
Db 1 CysArgArgCysAlaValIleValGlyAsnSerGlyAsnLeuArgGlySerGlyTyrglyGln 20
QY 454 GAGATAGATCGATCTCTCTCGATTGGAGATGACATGCCCCACCAAGGTTATGGA 513
Db 21 GluValAspSerHisAsnPhelleMetArgMetAsnGlnAlaProThrValGlyPheGlu 40
QY 514 GAAGATGTCGCGCATGACCATGATTCAGTGTGTGTCCTCATACCAGCGTCTCTTTG 573
Db 41 LysAspValGlySer-----ArgThrHisHisPheMetTyProGluSer 56
QY 574 CTAATA-----AACCTGATATTATTTTCAAGGAAGCGAAT 609
Db 57 AlaLysGlnIleTyraAsnProAlaPhePheCysAspGluValAsn 71
RESULT 9
US-09-334-601-21
; Sequence 21, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Kapitonov, Dmitri
; APPLICANT: Yu, Robert
; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Rattus gen. sp.
US-09-334-601-21
Alignment Scores:
Pred. No.: 3,28e-07 Length: 90
Score: 141.50 Matches: 32
Percent Similarity: 52.00% Conservative: 7
Best Local Similarity: 42.67% Mismatches: 29

```

Query Match:      5.94%      Indels:      7
DB:                3          Gaps:        2

US-09-714-936-218 (1-1294) x US-09-334-601-21 (1-90)

QY 394 TGTGACCTTTGTGCGCAGTGTCAAACTCAGCTCAGATGGTGGCCAGAGGTGGGAAT 453
Db 119 ValValValGlyAsnGlyGlyValLeuLysAsnLysThrLeuGlyAlaThrIleAspSer 138

QY 466 TCCCTCGCATTTGGGAGAAATGAACAATGCCCCACCAAGAGTTATGAAGAAGATGTCGCG 525
Db 139 TyrAspValIleleargMetAsnGlyProValLeuGlyHisGluGluGluValGly 158

QY 526 CGCATGACCATGATTGGATTTGTCTCCCATACAGCGTTCTCTTTGCTAAAAACCCCT 585
Db 159 ThrArgThrThrPheArg-----LeuPheTyrPro 168

QY 586 GATTATTTTTCAGGAAGCGGAAT-----ACTACTATTGTGTATT 627
Db 169 GluSerValPheAspSerSerHisTyrAspProAsnThrThrAlaValLeuValVal 188

QY 628 TGGGACCTTTTCGCAATATAGGAGAAATGCGCATGGCATGCTTACACATGTTGAA 687
Db 199 PheLysPro-----GluAspLeuArgTyrLeuValGluLeuLeuLeuGly 203

QY 688 AAGACACTTGGT-----AATCATCCGAAT 711
Db 204 LysLysIleAsnThrGlnGlyPheTyrPheThrProAlaLeuLysLeuIleTyrLysGln 223

QY 712 GCCCAATA-----TAGTGACCAACAGAGAAG-----CGCATGAGTTAC 750
Db 224 TyrGlnIleargIleLeuAspProTyrIleThrSerGluAlaAlaPheGlnMetLeuArg 243

QY 751 TGTGATGAGCTTTTAAAGAA 774
Db 244 PheProArgValPheProLysasp 251

RESULT 11
US-09-334-601-22
; Sequence 22, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Yu, Robert
; APPLICANT: Kapitonov, Dmitri
; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 22
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-334-601-22

Alignment Scores:
Pred. No.:      7.09e-07      Length:      90
Score:          138.50      Matches:      31
Percent Similarity: 52.00%      Conservative: 8
Best Local Similarity: 41.33%      Mismatches: 29
Query Match:    5.81%      Indels:      7
DB:            3          Gaps:        2

US-09-714-936-218 (1-1294) x US-09-334-601-22 (1-90)

QY 394 TGTGACCTTTGTGCGCAGTGTCAAACTCAGCTCAGTGTGGCCAGAGGTGGGAAT 453
Db 1 CysArgArgCysAlaValValGlyAsnSerGlyAsnLeuArgGlySerGlyTyrGln 20

QY 454 GAGATAGATCGATCCCTCGCATTTGGAGAAATGAACAATGCCCCACCAAGAGTTATGA 513
Db 21 GluValAspSerHisAsnPheIleMetArgMetAsnGlnAlaProThrValGlyPheGlu 40

QY 514 GAAGATCTCGCCCGCATGACCATGATTGCTGCTCCATACAGCGTTCTCTTTG 573
Db 41 LysAspValGlySer-----ArgThrThrHisPheMetTyrProGluSer 56

QY 574 CTAATA-----AACCTGATATTATTTTTCAGGAAGCGGAAT 609
Db 57 AlaLysGlnIleTyrAsnProAlaPhePheCysAspGluValAsn 71

RESULT 10
US-09-334-601-10
; Sequence 10, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Yu, Robert
; APPLICANT: Kapitonov, Dmitri
; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-334-601-10

Alignment Scores:
Pred. No.:      8.39e-07      Length:      329
Score:          140.50      Matches:      61
Percent Similarity: 35.82%      Conservative: 35
Best Local Similarity: 22.76%      Mismatches: 77
Query Match:    5.90%      Indels:      95
DB:            3          Gaps:        11

US-09-714-936-218 (1-1294) x US-09-334-601-10 (1-329)

QY 193 GTGATGCTGTGAGCTTCATAGCAGCGTTCTCTTTCTCTGCTGGTT----- 237
Db 5 LeuValAlaIlePheLeuSerSerIlePheLeuTyrTyrValLeuTyrCysIleLeuTyr 24

QY 238 -----GTGCGTCTTGTAAATGAATGAATGATTC 264
Db 25 GlyThrAsnGlyTyrTrpPheProAlaGluGluMetArgThrArgAsnValAsn--- 43

QY 265 CCATTCTACTAAACTGCTTTGGCAACCT----- 294
Db 44 -----AsnCysPheLysLysProAlaPheAlaAsnLeuLeuArgPheProGln 59

QY 295 -----GGTACAAG 303
Db 60 LeuTyrProPheLeuCysArgAlaAspPheIleLysValAlaAlaMetSerGlyThrAsn 79

QY 304 TGGATACCATCTCTCATACATACAGCGGCCCTTCGACTCATGATGATACATAAT 363
Db 80 AsnPheProLeuProTyrGlyIle---LysThrPheGluThrTyrPheSerSerAlaLeu 98

QY 364 GTGAGACACAA-----GAGCCTTTGCACTGGACCTGTGACCTTTGT 405
Db 99 SerLysLeuGlnSerCysAspLeuPheAspGluPheAspArgValProCysLysArgCys 118

```

```

QY 406 GCCATAGTGTCAAACCTCAGTGTGGCCAGAGGTGGGAATGAGATGATGCGA 465
Db 119 ValValValGlyAsnGlyGlyValLeuLysAsnLysThrLeuGlyAlaThrIleAspSer 138

QY 466 TCCCTCGCATTTGGGAGAAATGAACAATGCCCCACCAAGAGTTATGAAGAAGATGTCGCG 525
Db 139 TyrAspValIleleargMetAsnGlyProValLeuGlyHisGluGluGluValGly 158

QY 526 CGCATGACCATGATTGGATTTGTCTCCCATACAGCGTTCTCTTTGCTAAAAACCCCT 585
Db 159 ThrArgThrThrPheArg-----LeuPheTyrPro 168

QY 586 GATTATTTTTCAGGAAGCGGAAT-----ACTACTATTGTGTATT 627
Db 169 GluSerValPheAspSerSerHisTyrAspProAsnThrThrAlaValLeuValVal 188

QY 628 TGGGACCTTTTCGCAATATAGGAGAAATGCGCATGGCATGCTTACACATGTTGAA 687
Db 199 PheLysPro-----GluAspLeuArgTyrLeuValGluLeuLeuLeuGly 203

QY 688 AAGACACTTGGT-----AATCATCCGAAT 711
Db 204 LysLysIleAsnThrGlnGlyPheTyrPheThrProAlaLeuLysLeuIleTyrLysGln 223

QY 712 GCCCAATA-----TAGTGACCAACAGAGAAG-----CGCATGAGTTAC 750
Db 224 TyrGlnIleargIleLeuAspProTyrIleThrSerGluAlaAlaPheGlnMetLeuArg 243

QY 751 TGTGATGAGCTTTTAAAGAA 774
Db 244 PheProArgValPheProLysasp 251

RESULT 11
US-09-334-601-22
; Sequence 22, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Yu, Robert
; APPLICANT: Kapitonov, Dmitri
; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 22
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-334-601-22

Alignment Scores:
Pred. No.:      7.09e-07      Length:      90
Score:          138.50      Matches:      31
Percent Similarity: 52.00%      Conservative: 8
Best Local Similarity: 41.33%      Mismatches: 29
Query Match:    5.81%      Indels:      7
DB:            3          Gaps:        2

US-09-714-936-218 (1-1294) x US-09-334-601-22 (1-90)

QY 394 TGTGACCTTTGTGCGCAGTGTCAAACTCAGCTCAGTGTGGCCAGAGGTGGGAAT 453
Db 1 CysArgArgCysAlaValValGlyAsnSerGlyAsnLeuArgGlySerGlyTyrGln 20

QY 454 GAGATAGATCGATCCCTCGCATTTGGAGAAATGAACAATGCCCCACCAAGAGTTATGA 513
Db 21 AspValAspGlyHisAsnPheIleMetArgMetAsnGlnAlaProThrValGlyPheGlu 40

QY 514 GAAGATCTCGCCCGCATGACCATGATTGCTGCTCCATACAGCGTTCTCTTTG 573
Db 41 GlnAspValGlySer-----ArgThrThrHisPheMetTyrProGluSer 56

```

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

```

Query Match:      5.77%      Indels:      45
DB:               2          Gaps:        4
US-09-714-936-218 (1-1294) x US-08-446-875-10 (1-375)
QY 283 TTGGACAACCTGGTACAAAGTGGATACCATCTCTACACATACAGCGGCCCTTCGA 342
Db 102 PheSerLysProAlaProMetPheLeuAspSerPheArgLysTrpAlaArgIleArg 121
QY 343 ACTCAGTATGATACATAAATGTGAAGACACAGAGCCTTTG----- 384
Db 122 GluPheValProPheGlyIleLysGlyGlnAspAsnLeuLysAlaIleLeuSer 141
QY 385 -----CAACTGGACTGTGACCTTTGGCC 408
Db 142 ValThrLysGluThrArgLeuThrProAlaLeuAspSerLeuArgCysArgCysIle 161
QY 409 ATAGTGTCAAACTCAGTTCAGATGGTGGCCAGAAAGTGGAAATCAGATAGATCGATCC 468
Db 162 IleValGlyAsnGlyGlyValLeuAlaAsnLysSerLeuGlySerArgIleAspTyr 181
QY 469 TCCTGCATTGGAGAAATGAACAATGCCCCACCAAGGTTATGAAGAAGATGTGGCCGC 528
Db 182 AspIleValValArgLeuAsnSerAlaProValLysGlyPheGluLysAspValGlySer 201
QY 529 ATGACCATGATTCGAGTTGTGTGCCATACACAGCGTTCCTCTTTGTCTAAACACCCCTGAT 588
Db 202 LysThrThrLeuArgIle-----ThrTyrProGluGlyAlaMetGlnArgProGlu 218
QY 589 TATTTTTCAGGAAGCAGGAATCTACTATTGTGTTATT---TGGGACCTTTCGCAAT 645
Db 219 GlnTyrGluArgAspSerLeuPheValLeuAlaGlyPheLysTrpGlnAspPheLysTrp 238
QY 646 ATGAGGAAGATGCAATGCAATGCTGTTACACACATGTTGAAAGACACATGCGTATCTAT 705
Db 239 LeuLys----- 240
QY 706 CGGAATGCCAAATATACGTGACCAAGAGCAGATGAGTTACTGTGATGAGTTT 765
Db 241 -----TyrIleValTyrLysGluArgValSerAlaSerAspGlyPheTrp 255
QY 766 AAG 768
Db 256 Lys 256

RESULT 14
US-08-102-385G-10
; Sequence 10, Application US/08102385G
; Patent No. 5962294
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, Serge
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzhiradzky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, 38th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/102,385G
; FILING DATE: 04-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925369
; FILING DATE: 04-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oigenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 97-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310)788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-102-385G-10
Alignment Scores:
Pred. No.: 1.94e-06 Length: 375
Score: 137.50 Matches: 39
Percent Similarity: 40.33% Conservative: 34
Best Local Similarity: 21.55% Mismatches: 63
Query Match: 5.77% Indels: 45
DB: 2 Gaps: 4
US-09-714-936-218 (1-1294) x US-08-102-385G-10 (1-375)
QY 283 TTGGACAACCTGGTACAAAGTGGATACCATCTCTACACATACAGCGGCCCTTCGA 342
Db 102 PheSerLysProAlaProMetPheLeuAspSerPheArgLysTrpAlaArgIleArg 121
QY 343 ACTCAGTATGATACATAAATGTGAAGACACAGAGCCTTTG----- 384
Db 122 GluPheValProPheGlyIleLysGlyGlnAspAsnLeuLysAlaIleLeuSer 141
QY 385 -----CAACTGGACTGTGACCTTTGGCC 408
Db 142 ValThrLysGluThrArgLeuThrProAlaLeuAspSerLeuArgCysArgCysIle 161
QY 409 ATAGTGTCAAACTCAGTTCAGATGGTGGCCAGAAAGTGGAAATCAGATAGATCGATCC 468
Db 162 IleValGlyAsnGlyGlyValLeuAlaAsnLysSerLeuGlySerArgIleAspTyr 181
QY 469 TCCTGCATTGGAGAAATGAACAATGCCCCACCAAGGTTATGAAGAAGATGTGGCCGC 528
Db 182 AspIleValValArgLeuAsnSerAlaProValLysGlyPheGluLysAspValGlySer 201
QY 529 ATGACCATGATTCGAGTTGTGTGCCATACACAGCGTTCCTCTTTGTCTAAACACCCCTGAT 588
Db 202 LysThrThrLeuArgIle-----ThrTyrProGluGlyAlaMetGlnArgProGlu 218
QY 589 TATTTTTCAGGAAGCAGGAATCTACTATTGTGTTATT---TGGGACCTTTCGCAAT 645
Db 219 GlnTyrGluArgAspSerLeuPheValLeuAlaGlyPheLysTrpGlnAspPheLysTrp 238
QY 646 ATGAGGAAGATGCAATGCAATGCTGTTACACACATGTTGAAAGACACATGCGTATCTAT 705
Db 239 LeuLys----- 240
QY 706 CGGAATGCCAAATATACGTGACCAAGAGCAGATGAGTTACTGTGATGAGTTT 765
Db 241 -----TyrIleValTyrLysGluArgValSerAlaSerAspGlyPheTrp 255
QY 766 AAG 768
Db 256 Lys 256

RESULT 15
US-08-446-875-16

```

```

; Sequence 16, Application US/08446875
; Patent No. 5858751
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, Sorge
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzhiradzky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; IDENTIFICATION AND SYNTHESIS OF SIALYLTRANSFERASES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446, 875
; FILING DATE: July 12, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/102,385
; FILING DATE: August 4, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 111-197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-875-16

```

```

Alignment Scores:
Pred. No.: 3.09e-06 Length: 340
Score: 135.50 Matches: 60
Percent Similarity: 35.47% Conservative: 34
Best Local Similarity: 22.64% Mismatches: 76
Query Match: 5.69% Indels: 95
DB: 2 Gaps: 10

```

US-09-714-936-218 (1-1294) x US-08-446-875-16 (1-340)

```

QY 178 CTGAGAGAAAGTCGTGATGCTGAGCTTCATGACGCGTTCCTTTTCG----- 231
Db 4 LeuArgLysArgThrLeuLysValValThrPheLeuValLeuPheLeuThrSer 23
QY 232 -----CTGGTT 237
Db 24 PhePheLeuAsnTyrSerHisThrMetValAlaThrThrPheProLysGlnMetVal 43
QY 238 GTGGGCTGTGTAAGTGAAGTGAATTCCTGCTACTA----- 276
Db 44 LeuGluLeu-----SerGluAsnLeuLysArgLeuLeuLysHisArgProCysThrCys 61
QY 277 ---AACTGCTTTGGACACCTGGT---ACAAAGTGGATACCATTCCTCTAC----- 321
Db 62 ThrHisCysIleGlyGlnArgLysLeuSerAlaThrPheAspGluArgPheAsnGlnThr 81

```

```

QY 322 -----ACATACAGGGGGCCCC 336
Db 82 MetGlnProLeuLeuThrAlaGlnAsnAlaLeuLeuGluAsnAspThrTyrArgTrpTrp 101
QY 337 CTTGGAACCTCACTATGGATACATAAATGTGAAGACACACAGAGCCTTTGCACTGGAC--- 393
Db 102 LeuArgLeuGln-----ArgGluLysLysProAsnAsnLeuAsnAsp 115
QY 393 ----- 393
Db 116 ThrIleLysGluLeuPheArgValValProGlyAsnValAspProMetLeuGluLysArg 135
QY 394 -----TGTGACCTTTGTGCCATAGTCTCAAACTCAGTCAGATGGTGGCCAGAG 444
Db 136 SerValGlyCysArgArgCysAlaValValGlyAsnSerGlyAsnLeuArgGluSerSer 155
QY 445 GTGGGAATGAGATGATCGATCCTCTGCTATGGAATGAACAATGCCGCCACCAA 504
Db 156 TyrGlyProGluLysAspSerHisAspPheValLeuArgMetAsnLysAlaProThrAla 175
QY 505 GGTATGAGAAGATGTCGGCCGCGCATGACCATGATTCGAGTGTGCCATACCGCGTT 564
Db 176 GlyPheGluAlaAspValGlyThrLysThrHisHisLeuValTyrProGluSer--- 194
QY 565 CCTCTTTTGCTAAAAAACCCCTGATTATTTTTCAGGAGCGGATGACTACTATTGTT 624
Db 195 -----PheArgGluLeuGlyAspAsnValSerMet 204
QY 625 ATTTGGGACCTTTCCGCAATATGAGGAAGATGCAATGCGTCGTTTACACATGTTG 684
Db 205 IleLeuValProPheLysThrIle-----AspLeuGluTyrValValSerAlaIleThr 222
QY 685 AAAAGACAGTTGGT-----ATCTATCCGATGCCCAATATACGTGACCCAGAG 735
Db 223 ThrGlyThrIleSerHisThrTyrIleProValProAlaLysIleArgValLysGlnAsp 242
QY 736 AAGCGCATGAGTTAC 750
Db 243 LysIleLeuIleTyr 247

```

Search completed: September 12, 2003, 16:38:35
Job time : 31 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 15:04:59 ; Search time 60.5 Seconds
(without alignments)
6789.824 Million cell updates/sec

Title: US-09-714-936-218
Perfect score: 2383
Sequence: 1 ccggaattccgggtcgagc.....ttttctctctttttttg 1294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09714936/runat_10092003_151457_5670/app_query.fasta_1.1479
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=h2p.rag -MINMATCH=0.1 -LOOFC=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714936_ECGW_1.64 -runat_10092003_151457_5670 -NCP=6 -ICPU=3
-NO_MAP -LARGESQUERY -NBS_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1273	53.4	363	23	AAE21617 Human gene 1 encod
2	1177	49.4	221	23	ABP64885 Human protein SEQ
3	1100	46.2	305	23	ABG65165 Human albumin fusi
4	1100	46.2	305	23	ABG65166 Human albumin fusi
5	1100	46.2	305	23	AAE21578 Human gene 1 encod
6	1100	46.2	305	23	AAE21595 Human gene 1 encod
7	1095	46.0	210	22	AAU29291 Human PRO polypept
8	1095	46.0	210	22	AAE05186 Human drug metabol
9	1095	46.0	210	24	ABU71379 Human PRO1777 prot
10	1095	46.0	210	24	ABU65836 Human secreted/tra
11	1095	46.0	210	24	ABU66169 Novel human secret
12	1095	46.0	210	24	ABU67673 Human secreted/tra
13	1095	46.0	210	24	ABU65531 Human PRO polypept
14	1095	46.0	210	24	ABU58667 Human PRO polypept
15	1095	46.0	210	24	ABU56203 Human secreted/tra
16	1095	46.0	210	24	ABU57198 Human PRO polypept
17	1095	46.0	210	24	ABU10777 Human secreted/tra
18	1074	45.1	210	21	AB422629 Human ORFX ORF2033
19	859	36.0	166	22	AAU23360 Novel human enzyme
20	703	29.5	136	23	ABG65163 Human albumin fusi
21	703	29.5	136	23	AAE21596 Human gene 1 encod
22	391.5	16.4	345	21	AAE24495 Human secreted pro
23	382	16.0	302	21	AAE28674 Human carbohydrate
24	382	16.0	302	21	AAE25764 Human secreted pro
25	382	16.0	302	22	AAE61614 Human protein HP03
26	382	16.0	302	22	AAE75350 Human secreted pro
27	382	16.0	303	21	AAE24441 Human secreted pro
28	328.5	13.8	336	22	AAE10705 Human sialyltransf
29	328.5	13.8	336	23	ABG96367 Human ovarian canc
30	328.5	13.8	336	23	AAE22149 Human TRNFR-11 pro
31	319.5	13.4	275	22	AAW41882 Human polypeptide
32	317.5	13.3	333	22	AAK39358 Human polypeptide
33	317.5	13.3	353	21	AAE41983 Human ORFX ORF1647
34	305	12.8	256	22	AAW41144 Human polypeptide
35	296	12.4	246	22	AAW40096 Human polypeptide
36	293.5	12.3	299	22	AAU99356 Human PRO1359 (UNQ
37	293.5	12.3	299	22	AAU29126 Human PRO polypept
38	293.5	12.3	299	22	AAU94771 Human protein sequ
39	293.5	12.3	299	22	AAE66105 Protein of the inv
40	293.5	12.3	299	24	ABU71214 Human PRO1359 prot
41	293.5	12.3	299	24	ABU65671 Human secreted/tra
42	293.5	12.3	299	24	ABU66004 Novel human secret
43	293.5	12.3	299	24	ABU67508 Human secreted/tra
44	293.5	12.3	299	24	ABU65366 Human PRO polypept
45	293.5	12.3	299	24	ABU58502 Human PRO polypept

ALIGNMENTS

RESULT 1
AAE21617
ID AAE21617 standard; Protein; 363 AA.
XX AAE21617;
AC
XX
DT 16-JUL-2002 (first entry)
DE Human gene 1 encoded secreted protein, SEQ ID NO:89.
KW Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
KW gene therapy.

XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABQ99471.
XX
XX
PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity -
XX
XX
PS Claim 20; SEQ ID 545; 394pp; English.
XX
XX
CC The present invention relates to novel human coding sequences
CC (ABQ9268-ABQ93608) and proteins (ABP64682-ABP65022). The sequences are
CC useful in therapeutic, diagnostic and research methods. The
CC polynucleotides may be used in the field of molecular biology as
CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
CC for the recombinant production of protein, or in generation of anti-sense
CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
CC sequence tags (ESTs) for identifying expressed genes or for physical
CC mapping of the human genome. The proteins may be used as molecular weight
CC markers, or as nutritional sources or supplements. The proteins may be
CC used to maintain and expand cell population in a totipotential or
CC pluripotential state useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
CC development of bio-sensors. The polynucleotides and proteins are useful
CC for preventing, treating or ameliorating disorders involving aberrant
CC protein expression or biological activity, e.g. haematopoietic disorders,
CC central/peripheral nervous system diseases, mechanical and traumatic
CC disorders, non-healing wounds, immune deficiencies and disorders,
CC infectious diseases caused by viral, bacterial or fungal infection,
CC autoimmune disorders, allergic reactions and conditions, coagulation
CC disorders, or cancer. The polynucleotide sequences of the invention were
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
CC in some cases, sequences obtained from one or more public databases.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 221 AA;

Alignment Scores:
Pred. No.: 4,17e-121 Length: 221
Score: 1177.00 Matches: 221
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.39% Indels: 0
DB: 23 Gaps: 0

US-09-714-936-218 (1-1294) x ABP64885 (1-221)

QY 166 ATGGCTGATCTCTGAGAGAAAGTCTGTGATGCTGTGAGCTTCATAGAGCGTTCCTT 225
DB 1 MetAlaCysIleLeuIysArgIysSerValIleAlaValSerPheIleAlaAlaPheLeu 20

QY 226 TTCTGCTGCTGTGGCTGCTGTAATAAGTCAAGTGAATTCCTCATCTACTAACTGCTTT 285
DB 21 PheLeuLeuValValIargLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40

QY 286 GGACACCTGGTACAAAGTGGATACCAFTCTCTACACATACAGCGCGCCCTTCGAACCT 345
DB 41 GlyGlnProGlyThrIysrPileProPheSerTyThrTyArgArgProLeuArgThr 60

QY 346 CACTATGGATACATAAATCTGAAGACACAGAGCGCTTGCACACTGGACTGTGACCTTTGT 405
DB 61 HisTyGlyTyIleAsnValIysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80

QY 406 GCCATAGTCAAACTCAGTTCAGTGTGGTGGCCAGCAAGGTGGGAATCAGATAGATCGA 465
DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnIysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTTGGAGATGAACAATGCCCCACCACAAAGGTTATGAAGAAGATGTCGGC 525

DB 101 SerSerCysIleTrpArgMetAsnAlaProThrIysGlyTyrgluGluAspValGly 120
QY 526 CGCATGACCATGATTCGAGTTGTCTCCCATACACAGCGTTCTCTTTTGTCTAAACCCCT 585
DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATPACTACTATTGTGTATTGGGACCTTCCGCAAT 645
DB 141 AspTyPhePheIysGluAlaAsnThrThrIleCysValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGGAAGATCGCAATGGCATCGTTTACACATGTTGAAAAAGACAGATTGTATCTAT 705
DB 161 MetArgIysAspGlyAsnGlyIleValTyAsnMetLeuLysIleThrValGlyIleTy 180
QY 706 CCGAATGCCCAATATACGTGACACACAGAGCGCATGAGTACTGTGATGAGATTTT 765
DB 181 ProAsnAlaGlnIleTyValThrThrGluLysArgMetSerTyCysAspGlyValPhe 200
QY 766 AAGAGGAAGACTGGGAAGACAGAGGGGCATGCAAGCGGACTGCTGATTCTACAGACT 825
DB 201 LysIysGluThrGlyLysAspArgGlyHisAlaArgLeuLeuSerThrAspThr 220
QY 826 TTT 828
DB 221 Phe 221

RESULT 3
ABG65165
ID ABG65165 standard; Protein; 305 AA.
XX
AC ABG65165;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #1840.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antifertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US11988.
XX
PR 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2002-010886/01.
XX
PT New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
XX
PS Claim 1; Page 1796-1797; 2102pp; English.
XX
CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or

CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
 CC fusion proteins of the invention.

XX Sequence 305 AA;

Alignment Scores:
 Pred. No.: 1.68e-112 Length: 305
 Score: 1100.00 Matches: 207
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 46.16% Indels: 0
 DB: 23 Gaps: 0

US-09-714-936-218 (1-1294) x ABG65165 (1-305)

QY 166 ATGGCGCTGCATCTCGAAGAAAGTCTGTGATTGCTGAGCTTCATAGACGCTTCCTT 225
 |||||
 DB 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20

QY 226 TTCCTGCTGGTGTGCTCTTGTAAATGAAGTAAATTCCTACTAAACGCTTT 285
 |||||
 DB 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40

QY 286 GGACACCTGTGTACAAAGTGTACCATCTCTACATACACAGCGCCCTTCGACCT 345
 |||||
 DB 41 GlyGlnProGlyThrLysTrpIleProPheSerIleThrIleArgArgProLeuArgIle 60

QY 346 CACTATGATACATAAATGTGAAGACACAGAGCGCTTGCACATGGAGCTGTGACCTTGT 405
 |||||
 DB 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80

QY 406 GCCATAGTGTCAAACTCAGTCAGATGTTGCCACAGAGTGGGAAATGAGATACATCGA 465
 |||||
 DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100

QY 466 TCCTCTGCTGATTTGGAGAATGAACAATGCCCCACCAAGGTATGAAGAAGATGTCGGC 525
 |||||
 DB 101 SerSerCysIleTyrPargMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120

QY 526 CGCATGACCATGATTCAGATTGTGTCCTATACACAGCTTCTCTTTTGTCTAAACCCCT 585
 |||||
 DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140

QY 586 GATTATTTTTCAGGAAGCGAATACATCTACTATTGTTATTTGGGACCTTCCGCAAT 645
 |||||
 DB 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160

QY 646 ATGAGAAAGATGCGCATCGCTGTTTACACATGTTTGAAGAAAGACAGTGTGATCAT 705
 |||||
 DB 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180

QY 706 CCCAATGCCAAATATACGTGCACACAGAGAGCCCATGAGTACTGTGATGAGGTTTTT 765
 |||||
 DB 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSertyrCysAspGlyValPhe 200

QY 766 AAGAAGAAACTGGGAAGCAGG 789
 |||||
 DB 201 LysLysGluThrGlyLysAspArg 208

RESULT 4
 ABG65166
 ID ABG65166 standard; Protein; 305 AA.
 XX
 AC ABG65166;

XX 27-AUG-2002 (first entry)
 DT Human albumin fusion protein #1841.
 XX
 DE
 XX
 KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antifertility; antiinflammatory; antiulcer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX WO200177137-A1.
 FN 18-OCT-2001.
 PD 12-APR-2001; 2001WO-US11988.
 XX 12-APR-2000; 2000US-229358P.
 XX 25-APR-2000; 2000US-199384P.
 XX 21-DEC-2000; 2000US-256931P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Haseltine WA;
 XX WPI; 2002-010886/01.
 DR
 XX
 XX New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein -
 XX
 PS Claim 1; Page 1797-1798; 2102pp; English.
 XX
 CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
 CC fusion proteins of the invention.

XX Sequence 305 AA;

Alignment Scores:
 Pred. No.: 1.68e-112 Length: 305
 Score: 1100.00 Matches: 207
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 46.16% Indels: 0
 DB: 23 Gaps: 0

US-09-714-936-218 (1-1294) x ABG65166 (1-305)

QY 166 ATGGCGCTGCATCTCGAAGAAAGTCTGTGATTGCTGAGCTTCATAGACGCTTCCTT 225
 |||||
 DB 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20

QY 226 TTCCTGCTGGTGTGCTCTTGTAAATGAAGTAAATTCCTACTAAACGCTTT 285
 |||||
 DB 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40

Db 101 SerSerCysIleTTPArgMetaAsnAlaProThrLysGlyTyrgluAspValGly 120
QY 526 CCGATGACACGATTCAGTTGTCGCCATACCGAGCTTCCTCTTTGCTAAACACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTTTATTGGGACCTTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTyrPheLysAsnPro 160
QY 646 ATGAGGAAGATGCAATGCGATCGTTTACACATGTGTGAAAAGACAGTGTGATCAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGATGCCCAATATACGTGACACAGAGAGCGCATGATTACGTGATGGAGTTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGGAACCTGGAGGACAGG 789
Db 201 LysLysGluThrGlyLysAspArg 208
RESULT 6
AAE21595
ID AAE21595 standard; Protein; 305 AA.
XX AC AAE21595;
XX AC
XX 16-JUL-2002 (first entry)
XX Human gene 1 encoded secreted protein H9PR39, SEQ ID NO:67.
DE Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
KW diabetes; infection; wound healing; vulnerability; chemotaxis; food additive;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..30
FT /label= Signal_peptide
FT 31..305
FT /note= "Human mature secreted protein"
XX
XX W0200222654-A1.
XX
XX 21-MAR-2002.
XX
XX 17-JAN-2001; 2001WO-US01385.
XX
XX 12-SEP-2000; 2000US-231969P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI Ni J;
XX
XX WPI; 2002-315684/35.
DR N-PSDB; AAD34123.
XX
XX Seventeen nucleic acid molecules encoding human secreted proteins,
PT useful for treating and preventing cancer, immune disorders (e.g.
PT Addison's disease, and allergies), and cardiovascular disorders (e.g.
PT myocardial ischemias).
XX
PS Claim 11; Page 450-451; 483pp; English.
XX

CC AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted
CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 17 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of cancer,
CC proliferative disorders, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney
CC disorders, gastrointestinal disorders, pregnancy-related disorders,
CC endocrine disorders, and infections. The proteins can also be used to aid
CC wound healing and epithelial cell proliferation, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, for supporting
CC cell culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties. The
CC present sequence represents a human secreted protein of the invention.
XX
SQ Sequence 305 AA;
Alignment Scores:
Pred. No.: 1.88e-112 Length: 305
Score: 1100.00 Matches: 207
Percent Similarity: 99.52% Conservatve: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 46.16% Indels: 0
DB: 23 Gaps: 0
US-09-714-936-218 (1-1294) x AAE21595 (1-305)
QY 166 ATGGCTGCATCTCTGAAGAAAGTCTGTGATTCGTGAGCTTCATAGACGCTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaIlePheLeu 20
QY 226 TTCCTGCTGGTGGCTCTTGTAAATGAAGTGAATTCCTCATTCCTGCTAAACGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 SGACACCTGCTCAAGTGGATACCATCTCCCTACACATACAGGGCGCCCTTCGACT 345
Db 41 GlyGlnProGlyThrLysIleProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTGAAGACACACAGAGCTTTCGACTGTGACCTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysLeuCys 80
QY 406 GCATAGTGCACACTCAGTCAGATGTTGCCAGAGGTGGGAAATGAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTGGAGATGACCAATGCCCCACCAAGGTTTATGAAGAGATGTCGGC 525
Db 101 SerSerCysIleIlePargMetAsnAlaProThrLysGlyTyrGluLeuAspValGly 120
QY 526 CGCATGACCATGATTCGAGTGTGTCCTACACGCTTCCTCTTTTGTAAAAACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTTTATTGGGACCTTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTyrPheLysAsnPro 160
QY 646 ATGAGGAAGATGCAATGCGATCGTTTACACATGTGTGAAAAGACAGTGTGATCAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180

QY 706 CCAGATGCCAATATACGTGACACAGAGCGCATGAGTACTGTGATGGAGTTT 765
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 131 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAGGAACACTGGGAGGACAGG 789
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 201 LysLysGluThrGlyLysAspArg 208

RESULT 7
 AAU29291
 ID AAU29291 standard; Protein; 210 AA.
 XX AC
 XX AAU29291;
 XX AC
 DT 18-DEC-2001 (first entry)
 XX
 DE Human PRO polypeptide sequence #268.
 XX
 KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200168848-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06520.
 XX
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186368P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000WO-US06684.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 30-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000WO-US1370S.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 05-JUN-2000; 2000US-209332P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30928.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 XX
 PA (GETH) GENENTECH INC.
 XX

PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2001-602746/68.
 DR N-PSDB; AAS46192.
 XX
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -
 XX
 PS Claim 11; Fig 536; 774pp; English.
 XX
 CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 210 AA;

Alignment Scores:

Pred. No.: 4,98e-112 Length: 210
 Score: 1095.00 Matches: 206
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 45.95% Indels: 0
 DB: 22 Gaps: 0

US-09-714-936-218 (1-1294) x AAU29291 (1-210)

QY 166 ATGGCTGTCATCTGAGAGAAAGTCTGTGATTGCTGTGATGCTCATAGACGCTTCCT 225
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
 QY 226 TTCCTGCTGGTGTGGCTTTGTAAATGAAGTGAATTTCCCATTCCTACTAACTGCTTT 285
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40
 QY 286 GGACAACTGGTACAAAGTGGATACCATTCCTCTACACATACAGCGGCCCTTCGAACT 345
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgProLeuArgThr 60
 QY 346 CACTATGGATACATAAATGTGAACACACAGAGCCTTGAACCTGGACTGTGACCTTTGT 405
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAsnCysAspLeuCys 80
 QY 406 GCCATAGTGTCAAACTCAGGTACAGTGGTGGCCAGAGGTGGGAAATGAGATAGATCGA 465
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
 QY 466 TCCTCTGCTGATTTGGAGATGAACAATGCCCCACCAAGAGTTATGAGAGATGTCGGC 525
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTyrGluGluAspValGly 120
 QY 536 CGCATGACCATGATTCGAGTTGTGTCCATACAGAGCTTCCTCTTTTCTGCTAAACCCCT 585
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
 QY 586 GATTATTTTTCAGGAGCGAATACTACTATTGTGTATTGGGGACCTTTCGCGAAT 645
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyPropheArgAsn 160

QY 646 ATGAGGAAGATGCAATGCGTCTGTTACACATGTTGAAAAAGACAGTGTGATCTAT 705
CC
Db 161 MetArgLysaspGlyasnGlylleValThrAsnMetLeuLysThrValGlylleYr 180
QY 706 CGGATGCCCCAATATACGTGACCCACAGAGAGCGCATGAGTTACTGTGAGGAGTTT 765
CC
Db 181 ProAsnAlaGlnlleYrValThrThrGluLysArgMetSerYrCysaspGlyValPhe 200
QY 766 AAGAGGAACTGGGAGGAC 786
CC
Db 201 LysLysGluThrGlyLysasp 207
CC
RESULT 8
AAE05186
ID AAE05186 standard; Protein; 210 AA.
XX
AC AAE05186;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human drug metabolising enzyme (DME-17) protein.
XX
KW Human; drug metabolising enzyme; DME-17; immunosuppressive; gene therapy;
KW cytostatic; autoimmune disorder; inflammatory disorder; atherosclerosis;
KW osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism;
KW rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia;
KW developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy;
KW thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma;
KW gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma;
KW actinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma;
KW cell proliferative disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= Signal_peptide
FT Protein 31..210
FT /note= "Mature drug metabolising enzyme (DME-17) protein"
FT Domain 80..157
FT /label= Lumenal_domain
FT /note= "Sialyltransferase"
XX
FN W0200151638-A2.
XX
PD 19-JUL-2001.
XX
PE 12-JAN-2001; 2001WO-US01174.
XX
PF 14-JAN-2000; 2000US-0176139.
PR 21-JAN-2000; 2000US-0177443.
PR 28-JAN-2000; 2000US-0178574.
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX
PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI Ring HZ, Hillman JR, Yue H, Azimzal Y, Yao MG, Gandhi AR;
PI Nguyen DB, Tang YT, Lal P, Bandman O;
XX
DR WPI; 2001-425874/45.
DR N-PSDB; AAD09952.
XX
PT Drug metabolising enzymes and encoding polynucleotides, useful for
PT diagnosing, treating and/or preventing autoimmune, inflammatory, cell
PT proliferative, developmental, endocrine, eye, metabolic, and
PT gastrointestinal disorders -
XX
PS Claim 1; Page 150; 133pp; English.
XX
CC The present sequence is human drug metabolising enzyme (DME-17) protein.
CC Human DME and its nucleic acid molecule are useful for the diagnosis,
CC treatment and prevention of disorders associated with increased or
CC decreased expression of DME. Examples of such disorders include,

CC autoimmune/inflammatory disorder such as acquired immune deficiency
CC syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative
CC disorder such as actinic keratosis, atherosclerosis; developmental
CC disorder such as epilepsy, anaemia; endocrine disorder such as
CC acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as
CC diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis;
CC metabolic disorder such as Addison's disease, obesity; gastrointestinal
CC disorder such as anorexia, dysphagia and hepatic tumours including
CC nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for
CC creating 'knockin' humanised animals (pigs) or transgenic animals (mice
CC or rats) to model human disease. DME DNA is also in useful is gene
CC therapy. DME and its immunogenic fragments are useful for screening
CC libraries of compounds in several drug screening assays.
XX
SQ Sequence 210 AA;
Alignment Scores:
Pred. No.: 4.98e-112 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 22 Gaps: 0
US-09-714-936-218 (1-1294) x AAE05186 (1-210)
QY 166 ATGGCTCCATCTCTGAGAGAAAGTCTGATTCTGTGAGCTTCATAGCAGCTTCCTT 225
CC
Db 1 MetAlaCyslleuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCTCTGCTGGTGTGGCTCTTGTAAATGAAGTGAATTCCTCATTTAACTGGCTTT 285
CC
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 SGAGAACTGGTCAAAAGTGGATACCATCTCTACATACAGAGCGCCCTTCGAACT 345
CC
Db 41 GlyGlnProGlyThrLysTrpIleProPheSerYrThrYrArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTAAGACACAGAGCGCTTGTGCACTGGACGTGACCTTGT 405
CC
Db 61 HisTyrGlyTrpIleAsnValLysThrGlnGluProLeuGlnLeuAspCysLeuCys 80
QY 406 GCATAGTGTCAAACTCAGTGCAGATGGTTGCCAGAAAGTGGGAAATGAGATATCGA 465
CC
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCCTGCATTGGAGATGAACAATGCCCCCAAGGTTATGAGAGATGTGGC 525
CC
Db 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTrpGluLysValGly 120
QY 526 CGCATGACCATGATCGAGTTGTGTCCTACCATACAGCGCTTCTCTTTGCTAAAAACCT 585
CC
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATATATTTTCAAGGAAGCAATCTACTATTTGTATTGTTGGGACCTTTCGCAAT 645
CC
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleYrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGGAAAGATGCAATGCGTCTGTTACACATGTTGAAAAGACAGTGTGATCTAT 705
CC
Db 161 MetArgLysaspGlyLysnGlylleValYrAsnMetLeuLysThrValGlylleYr 180
QY 706 CGGAATGCCAAATATACGTGACCCACAGAGAGCGCATGAGTTACTGTGAGGAGTTT 765
CC
Db 181 ProAsnAlaGlnlleYrValThrThrGluLysArgMetSerYrCysaspGlyValPhe 200
QY 766 AAGAGGAACTGGGAGGAC 786
CC
Db 201 LysLysGluThrGlyLysasp 207
CC
RESULT 9
ABU71379
ID ABU71379 standard; Protein; 210 AA.

[illegible]

```

PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090698P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091633P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-098349P.
PR 18-AUG-1998; 98US-098593P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Alignment Scores:
Pred. No.: 4,98e-112
Score: 210
Length: 210
Matches: 206
Conservative: 0
Percent Similarity: 99.52%
Best Local Similarity: 99.52%
Query Match: 45,95%
DB: 24
Gaps: 0

US-09-714-936-218 (1-1294) x ABU71379 (1-210)

QY 166 ATGCGCTGCATCCTCGAAGAAAGCTGTGATGTGAGCTTCATPAGCAGCGTTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGGTGTGCGTCTGTAAATGAAGTGAATTTCCCATTCGCTACTGAACGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACACCTGGTACAAAGTGGATACCATTCCTCTACACATACAGCGCGCCCTTCGAAC 345

```

```

Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTGAAGACACACAGAGCCTTTGCAACTGGACTGTGACCTTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
QY 406 GCATAGTGTCAAACTCAGGTGAGATGGTGGCCAGAGAGTGGGAATGAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTTGGGAATGAACAATGCCCCCAAGGTTTATGAAGAAGATGTCGGC 525
Db 101 SerSerCysIleIleTrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CCATGACCATGATGAGTGTGTCGCCATCCAGCAGCTTCCTCTTTTGTCTAAAACCCCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTCAAGGAAGCAATACTACTATTGTCTTTATTTGGGACCTTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGAAAGATGCAATGCGCATCGTTTACAACTGTGTAAGACACAGCTTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CCGATGCCCAATAATACGTGACACACAGAGAGCGCATGAGTACTCTGCTGAGGTTTTT 765
Db 181 ProAsnAlaGlnIleTyrValThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGAACTGGGAAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207

RESULT 10
ABU65836
ID ABU65836 standard; Protein; 210 AA.
AC ABU65836;
XX
XX 19-MAY-2003 (first entry)
DT
DE
DE Human secreted/transmembrane protein, SEQ ID 536.
KW Human; PRO; secreted protein; transmembrane protein;
KW cytosolic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
XX 20-FEB-2003.
PD
XX
XX 02-JUL-2002; 2002US-0188767.
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 02-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.

```


PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05844.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063670P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077642P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082566P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084633P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087755P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088126P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 04-JUL-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095598P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.

```
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Alignment Scores:
Pred. No.: 4.98e-112 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 24 Gaps: 0

US-09-714-936-218 (1-1294) x ABU5836 (1-210)
QY 166 ATGGCTGCATCCCTGAAGAAAGTCTGTGATTGCTGTGACCTCATACAGCGTTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGTTGCTGCTTGTAAATGAAGTAATTCCTGCTGCTGCTGCTGCTGCTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACTGCTGCAAAAGTGGATACATCTCTCTACATACAGCGGCGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysTrpIleProPheSerThrThrArgArgProLeuArgThr 60
QY 346 CACTATGATGATCAATAATGTAAGACACACAGAGCTTGGCACTGGACGTGACCTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysLeuCys 80
QY 406 GCATAGTGTCAAACTCAGTCAAGTGTGCTGCGCAGAGCTGGAAATCAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTGGAGATGAATGCAATGCCCCCAAGGTTATGAAGAAGATGTCGGC 525
Db 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CCATGACACATGATCAGTGTGTCCTCATACAGAGCTTCTCTTTTGTGTTAAACCCCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTCAAGGAAGCAATACTACTATTGTTGTTATTTGGGACCTTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleThrValIleThrIleThrPheArgAsn 160
QY 646 ATGAGAAGATGCAATGATCGTTTACACATGTTGAAAGACAGTTCGTATCAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleThr 180
QY 706 CCGATGCCCAATATACGTGACACACAGAGAGCGATGATGCTGATGAGTTCCTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGGAAGTGGGAAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207
```

```
RESULT 11
ABU56169
ID ABU56169 standard; Protein; 210 AA.
XX
XX AC ABU56169;
XX
XX DT 20-MAY-2003 (first entry)
XX
XX DE Novel human secreted and transmembrane protein PRO177.
XX
XX KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
XX OS Homo sapiens.
XX
XX PN US2003036157-A1.
XX
XX PD 20-FEB-2003.
XX
XX PF 02-JUL-2002; 2002US-0188769.
XX
XX PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 98WO-US05028.
PR 14-MAY-1999; 98WO-US10733.
PR 02-JUN-1999; 98WO-US12252.
PR 01-SEP-1999; 98WO-US20111.
PR 15-SEP-1999; 98WO-US21090.
PR 01-DEC-1999; 98WO-US28301.
PR 02-DEC-1999; 98WO-US28551.
PR 30-DEC-1999; 98WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US30952.
PR 08-NOV-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
```

PR	24-NOV-1997;	97US-066456P.	PR	17-JUN-1998;	98US-089653P.
PR	24-NOV-1997;	97US-066772P.	PR	18-JUN-1998;	98US-089908P.
PR	11-DEC-1997;	97US-069335P.	PR	19-JUN-1998;	98US-089952P.
PR	12-DEC-1997;	97US-069425P.	PR	22-JUN-1998;	98US-090246P.
PR	17-DEC-1997;	97US-069870P.	PR	22-JUN-1998;	98US-090252P.
PR	18-DEC-1997;	97US-068017P.	PR	22-JUN-1998;	98US-090254P.
PR	10-MAR-1998;	98US-077450P.	PR	24-JUN-1998;	98US-090429P.
PR	11-MAR-1998;	98US-077632P.	PR	24-JUN-1998;	98US-090435P.
PR	11-MAR-1998;	98US-077649P.	PR	24-JUN-1998;	98US-090444P.
PR	20-MAR-1998;	98US-078866P.	PR	24-JUN-1998;	98US-090461P.
PR	20-MAR-1998;	98US-078939P.	PR	24-JUN-1998;	98US-090535P.
PR	27-MAR-1998;	98US-079664P.	PR	24-JUN-1998;	98US-090540P.
PR	27-MAR-1998;	98US-079786P.	PR	25-JUN-1998;	98US-090676P.
PR	31-MAR-1998;	98US-080107P.	PR	25-JUN-1998;	98US-090678P.
PR	31-MAR-1998;	98US-080194P.	PR	25-JUN-1998;	98US-090678P.
PR	01-APR-1998;	98US-080327P.	PR	25-JUN-1998;	98US-090688P.
PR	01-APR-1998;	98US-080333P.	PR	25-JUN-1998;	98US-090690P.
PR	08-APR-1998;	98US-081049P.	PR	25-JUN-1998;	98US-090694P.
PR	08-APR-1998;	98US-081070P.	PR	25-JUN-1998;	98US-090695P.
PR	09-APR-1998;	98US-081195P.	PR	25-JUN-1998;	98US-090696P.
PR	15-APR-1998;	98US-081838P.	PR	26-JUN-1998;	98US-090662P.
PR	21-APR-1998;	98US-082568P.	PR	26-JUN-1998;	98US-090863P.
PR	21-APR-1998;	98US-082569P.	PR	26-JUN-1998;	98US-090863P.
PR	22-APR-1998;	98US-082704P.	PR	01-JUL-1998;	98US-091010P.
PR	22-APR-1998;	98US-082797P.	PR	01-JUL-1998;	98US-091359P.
PR	28-APR-1998;	98US-083322P.	PR	01-JUL-1998;	98US-091544P.
PR	29-APR-1998;	98US-083495P.	PR	02-JUL-1998;	98US-091478P.
PR	29-APR-1998;	98US-083496P.	PR	02-JUL-1998;	98US-091486P.
PR	29-APR-1998;	98US-083499P.	PR	02-JUL-1998;	98US-091626P.
PR	29-APR-1998;	98US-083559P.	PR	02-JUL-1998;	98US-091628P.
PR	05-MAY-1998;	98US-084366P.	PR	02-JUL-1998;	98US-091632P.
PR	06-MAY-1998;	98US-084414P.	PR	24-JUL-1998;	98US-094006P.
PR	07-MAY-1998;	98US-084639P.	PR	04-AUG-1998;	98US-095282P.
PR	07-MAY-1998;	98US-084640P.	PR	10-AUG-1998;	98US-095998P.
PR	07-MAY-1998;	98US-084643P.	PR	10-AUG-1998;	98US-096012P.
PR	15-MAY-1998;	98US-085579P.	PR	17-AUG-1998;	98US-096757P.
PR	15-MAY-1998;	98US-085580P.	PR	17-AUG-1998;	98US-096766P.
PR	15-MAY-1998;	98US-085582P.	PR	17-AUG-1998;	98US-096867P.
PR	15-MAY-1998;	98US-085700P.	PR	17-AUG-1998;	98US-096891P.
PR	18-MAY-1998;	98US-086023P.	PR	18-AUG-1998;	98US-096959P.
PR	22-MAY-1998;	98US-086392P.	PR	18-AUG-1998;	98US-097022P.
PR	22-MAY-1998;	98US-086486P.	PR	26-AUG-1998;	98US-097952P.
PR	28-MAY-1998;	98US-087098P.	PR	26-AUG-1998;	98US-097954P.
PR	28-MAY-1998;	98US-087208P.	PR	26-AUG-1998;	98US-097955P.
PR	02-JUN-1998;	98US-087609P.	PR	26-AUG-1998;	98US-097971P.
PR	02-JUN-1998;	98US-087759P.	PR	26-AUG-1998;	98US-097974P.
PR	03-JUN-1998;	98US-087827P.	PR	26-AUG-1998;	98US-098014P.
PR	04-JUN-1998;	98US-088025P.	PR	26-AUG-1998;	98US-098014P.
PR	04-JUN-1998;	98US-088028P.	PR	01-SEP-1998;	98US-098716P.
PR	04-JUN-1998;	98US-088029P.	PR	01-SEP-1998;	98US-098723P.
PR	04-JUN-1998;	98US-088033P.	PR	02-SEP-1998;	98US-098803P.
PR	05-JUN-1998;	98US-088326P.	PR	02-SEP-1998;	98US-098803P.
PR	05-JUN-1998;	98US-088167P.	PR	02-SEP-1998;	98US-098821P.
PR	05-JUN-1998;	98US-088202P.	PR	02-SEP-1998;	98US-098843P.
PR	05-JUN-1998;	98US-088212P.	PR	09-SEP-1998;	98US-099602P.
PR	05-JUN-1998;	98US-088217P.	PR	10-SEP-1998;	98US-099741P.
PR	09-JUN-1998;	98US-088655P.	PR	10-SEP-1998;	98US-099754P.
PR	10-JUN-1998;	98US-088722P.	PR	10-SEP-1998;	98US-099763P.
PR	10-JUN-1998;	98US-088738P.			
PR	10-JUN-1998;	98US-088740P.			
PR	10-JUN-1998;	98US-088811P.			
PR	10-JUN-1998;	98US-088824P.			
PR	10-JUN-1998;	98US-088825P.			
PR	10-JUN-1998;	98US-088826P.			
PR	11-JUN-1998;	98US-088861P.			
PR	11-JUN-1998;	98US-088863P.			
PR	11-JUN-1998;	98US-088876P.			
PR	12-JUN-1998;	98US-089090P.			
PR	12-JUN-1998;	98US-089105P.			
PR	16-JUN-1998;	98US-089512P.			
PR	16-JUN-1998;	98US-089514P.			
PR	17-JUN-1998;	98US-089538P.			
PR	17-JUN-1998;	98US-089598P.			
			Alignment Scores:		
			Pred. No.:	4.98e-112	210
			Score:	1095.00	Matches: 206
			Percent Similarity:	99.52%	Conservative: 0
			Best Local Similarity:	99.52%	Mismatches: 1
			Query Match:	45.95%	Indels: 0
			DB:	24	Gaps: 0
			US-09-714-936-218 (1-1294) x ABU66169 (1-210)		
QY	166	ATGGCTGCTCCTCAAGAGAACTCTGTGATTCGTGAGCTTCATAGAGCGTTCCTT	225		
Db	1	MetAlaCysIleLeuIysArgIysSerValIleAlaValSerPheIleAlaAlaPheLeu	20		
QY	226	TTCTCTGCTGTTGCTGCTTGTAAATGAAGTGAATTCCTCCACTGCTACTAACTGCTTT	285		

Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
 QY 286 GGACAACTGGTCAAAAGTGGATACCAATCTCTACACATACAGGCGCCCTTCGAAC 345
 Db 41 GlyGlnProGlyThrLysIrpLeProPheSerTyrThrTyrArgArgProLeuArgThr 60
 QY 345 CACATGGAACATTAATGTGAGACACAAAGAGCCCTTCAACTGGACTGTGACCTTGT 405
 Db 61 HistGlyTyrIleAsnValLysThrGlnGlnProLeuGlnLeuAsnCysPheLeuCys 80
 QY 406 GCATAGTGTCAAACTCAGGTGAGTGGTCCAGAAAGTGGGAAATCAGATAGTCGA 465
 Db 81 AlalIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAsnArg 100
 QY 466 TCCTCCGATTCAGGAATGAGAAATGCCCCACCAAGGTATGACAGAGATGTCGGC 525
 Db 101 SerSerCysIleIrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAsnValGly 120
 QY 526 CGCATGACCATGATTCAGTGTGTGTCCTCATCCATACAGCGTTCCTCTTTGCTAAAAACCT 585
 Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
 QY 586 GATTATTTTTCAGGAGCGAATACTACTATTGCTTATTGTTGTTGTTGTTGTTGTTGTTGTT 645
 Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleIrpGlyProPheArgAsn 160
 QY 646 ATGAGGAAGATGCAATGCGTTCGTTACACATGTTGAAAGACACAGTGTGATCTAT 705
 Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysThrValGlyIleTyr 180
 QY 706 CGCAATGCCAAATATACGTGACCCACAGAGAGCGCATGATGATGATGATGATGATGATGAT 765
 Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAGGAACTGGGAGGAC 786
 Db 201 LysLysGluThrGlyLysAsp 207
 RESULT 12
 ABU67673
 ID ABU67673 standard; Protein; 210 AA.
 XX AC ABU67673;
 XX DT 29-MAY-2003 (first entry)
 XX DE Human secreted/transmembrane protein (PRO) #268.
 KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
 KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
 KW tissue typing.
 OS Homo sapiens.
 XX PN US2003036162-AL.
 XX PD 20-FEB-2003.
 XX PF 12-JUL-2002; 2002US-0194423.
 PR 16-SEP-1998; 98WO-US19330.
 PR 07-OCT-1998; 98WO-US21141.
 PR 01-DEC-1998; 98WO-US25108.
 PR 08-MAR-1999; 99WO-US05028.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28551.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.

PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 29-AUG-2001; 2001WO-US27099.
 PR 26-JUN-1998; 98US-0105413.
 PR 07-OCT-1998; 98US-0168978.
 PR 06-NOV-1998; 98US-0187368.
 PR 07-DEC-1998; 98US-0202054.
 PR 03-MAR-1999; 99US-0254311.
 PR 14-MAY-1999; 99US-0311832.
 PR 14-MAY-1999; 99US-0380137.
 PR 25-AUG-1999; 99US-0380138.
 PR 25-AUG-1999; 99US-0380139.
 PR 25-AUG-1999; 99US-0380142.
 PR 18-OCT-1999; 99US-0403297.
 PR 12-NOV-1999; 99US-0423844.
 PR 22-AUG-2000; 2000US-0644848.
 PR 18-SEP-2000; 2000US-0664610.
 PR 18-SEP-2000; 2000US-0665350.
 PR 08-NOV-2000; 2000US-0709238.
 PR 20-DEC-2000; 2000US-0747259.
 PR 22-MAR-2001; 2001US-0816744.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 25-MAY-2001; 2001US-0866028.
 PR 05-JUN-2001; 2001US-0874503.
 PR 18-JUL-2001; 2001US-0908827.
 PR 30-JUL-2001; 2001US-0918585.
 PR 06-AUG-2001; 2001US-0924419.
 PR 13-AUG-2001; 2001US-0929404.
 PR 16-AUG-2001; 2001US-0931836.
 PR 28-AUG-2001; 2001US-0941992.
 PR 04-SEP-2001; 2001US-0946374.
 PR 15-JAN-2002; 2002US-0052586.
 (GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 Pan J, Smith V, Katanabe CK, Wood WI, Zhang Z;

WPI; 2003-332039/31.

N-PSDB; ACA05967.

New secreted and transmembrane PRO polypeptides and nucleic acids.
 useful in gene therapy, in chromosome and gene mapping, as chromosome
 markers, in tissue typing, and in chromosome identification -

Claim 11; Fig 536; 706pp; English.

The invention discloses human nucleic acids encoding secreted and
 transmembrane (PRO) polypeptides. Also disclosed is an antibody that
 specifically binds to the PRO polypeptide, a method for stimulating the
 release of tumour necrosis factor alpha (TNF-alpha) from human blood by
 contacting the blood a PRO polypeptide, a method for stimulating the
 proliferation or differentiation of chondrocyte cells by contacting the


```

PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081155P.
PR 13-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083435P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 05-MAY-1998; 98US-083559P.
PR 06-MAY-1998; 98US-084366P.
PR 07-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086352P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.

```

```

PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 01-JUL-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 02-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

```

Alignment Scores:

```

Pred. No.: 4,98e-112 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 24 Gaps: 0

```

US-09-714-936-218 (1-1294) x ABU65531 (1-210)

```

QY 166 ATGGCTGCATCTCGAGAGAAAGTCGTGCTGTGAGCTTCATAGCAGCGTTCCT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
QY 226 TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40
QY 286 GGACACCTGGTACAAAGTGGATACCATTCCTCTACATACAGCGGCCCTTCGAAC 345

```



```

PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 18-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090862P.
PR 25-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 26-JUN-1998; 98US-091019P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091638P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.

```

```

PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 01-SEP-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Alignment Scores:
Pred. No.: 4,98e-112 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 24 Gaps: 0

US-09-714-936-218 (1-1294) x ABUS8667 (1-210)

QY 166 ATGGCTGCATCCTGAAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCGTTCCTT 225
Db 1 MetalaCysIleLeuIysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGTTGTGGCTCTTCTTAATGAAGTGAATTCCTCATTCCTACTAACTGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACCTGGTACAAAGTGGATACCATCTCCTACACATACAGGGCCCTTCGCACT 345
Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGAAGACACAGAGCCTTTGCAACAGCTGTGACGCTTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
QY 406 GCCATAGTCAAACTCAGCTCAGATGTTGGCCAGAGTGGGAAATGAGATGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTTGGAGAATGAACAATGCCCCACCAAGGTTATGAAGAAGATGTGGC 525
Db 101 SerSerCysIleTrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 GCATGACCATGATTCGAGTTGTGCCATACAGGTTCCCTCTTTTCTAAAAACCCCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTTATTGTTGGGACCTTCGCGAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpIleProPheArgAsn 160
QY 646 ATGAGGAAAGATGCCAATGCCATCGTTTACAAACATGTTGAAAAAGACAGTTGGTATCAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysThrValGlyIleTyr 180
QY 706 CGGAATGCCCAATATACCTGACCACAGAGAGCGCATGAGTACTGTGTCGAGTTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAGGAAACTGGAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207

```


RESULT 15
ABU56203
ID ABU56203 standard; Protein; 210 AA.
XX
AC ABU56203;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted/transmembrane protein, PRO7177.
DE
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW antiarthritic; vulnery; tumour necrosis factor-alpha;
KW chondrocyte cell proliferation; chondrocyte cell differentiation;
KW tumour; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour;
KW bone disorder; cartilage disorder; arthritis; sports injury.
XX
OS Homo sapiens.
XX
PN US2003022298-A1.
XX
PD 30-JAN-2003.
XX
PF 20-JUN-2002; 2002US-0176913.
XX
PR 05-NOV-1997; 97WO-US20069.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 98WO-US00106.
PR 08-MAR-1999; 98WO-US05028.
PR 10-MAR-1999; 98WO-US05190.
PR 20-APR-1999; 98WO-US08615.
PR 14-MAY-1999; 98WO-US10733.
PR 02-JUN-1999; 98WO-US12252.
PR 01-SEP-1999; 98WO-US20111.
PR 08-SEP-1999; 98WO-US20594.
PR 13-SEP-1999; 98WO-US20944.
PR 15-SEP-1999; 98WO-US21090.
PR 15-SEP-1999; 98WO-US21547.
PR 05-OCT-1999; 98WO-US23089.
PR 29-NOV-1999; 98WO-US28214.
PR 30-NOV-1999; 98WO-US28311.
PR 30-NOV-1999; 98WO-US28409.
PR 01-DEC-1999; 98WO-US28301.
PR 02-DEC-1999; 98WO-US28634.
PR 02-DEC-1999; 98WO-US28551.
PR 02-DEC-1999; 98WO-US28564.
PR 16-DEC-1999; 98WO-US28565.
PR 20-DEC-1999; 98WO-US30095.
PR 20-DEC-1999; 98WO-US30911.
PR 20-DEC-1999; 98WO-US30999.
PR 22-DEC-1999; 98WO-US30720.
PR 30-DEC-1999; 98WO-US31243.
PR 30-DEC-1999; 98WO-US31271.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 22-MAR-2001; 2001WO-US09552.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078866P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.

PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.

Alignment Scores:
Pred. No.: 4 98e-112 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 24 Gaps: 0

US-09-714-936-218 (1-1294) x ASU56203 (1-210)

QY 166 ATGGCTGCATCTGAGAGAAAGTCTGTGATTCCTGTGAGCTTCATAGCAGCTTCTT 225
|||||
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCTCTGCTGTTGCTGCTTGTAAATGAAGTGAATTCCTCATTCCTACTAACTGCTTT 285
|||||

Db 21 PheLeuLeuValValArgLeuValValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GCACAACCTGGTACAAAAGTGGATACCAATTCTCTTACACATACAGCGCGCCCTTCGAACT 345
|||||
Db 41 GlyGlnProGlyThrLysrPileProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGGATACATAAATGTGAAGACACAGAGCCTTGCACACTGGACTGTGACCTTTGT 405
|||||
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGlnProLeuGlnLeuAsnCysAspLeuCys 80
QY 406 GCATAGTGTCAAACCTCAGTTCAGATGGTGGCCAGAAAGTGGGAAATGAGATAGATCGA 465
|||||
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCTGCTGATTTGGAGAAATGAACAATGCCCCCAAGGTTATGAAGAAGATGTGGC 525
|||||
Db 101 SerSerCysIleTrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CGCATGACCATGATTCGAGTTGTGTGCCATACAGCGTTCCTCTTTTCTTAAAAACCT 585
|||||
Db 121 ArgMetThrMetIleArgValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTATTGTGGGACCTTCCGCAAT 645
|||||
Db 141 AspTyrPhePheLysGluLysGlnLysThrThrIleTyrValIleTrpGlyPropheArgAsn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACATGTGTGAAAAGACACAGTTGGTATCTAT 705
|||||
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGAATGCCAAATATACGTGACCGACAGAGAGCGCATGAGTACTGTGTGAGGAGTTT 765
|||||
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGGAAACTGGGAAGGAC 786
|||||
Db 201 LysLysGluThrGlyLysasp 207

Search completed: September 12, 2003, 16:32:11
Job time : 76.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:24:09 ; Search time 87.5 Seconds
(without alignments)
7632.459 Million cell updates/sec

Title: US-09-714-936-218
Perfect score: 283
Sequence: 1 cgggaatttcgggtgcagc.....ttttttctttttttttt 1294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09714936/runat_10092003_151459_5692/app_query.fasta_1.1479
-DB=SPTREMBL_23 -QFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714936 -CGN_1_1_95 -runat_10092003_151459_5692 -NCPU=6 -ICPU=3
-NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1100	46.2	305	4	Q8NDV1	Q8ndv1 homo sapien

	2	1099	46.1	305	4	Q8N259
	3	411	17.2	234	13	Q9W6U6
	4	391	16.4	302	11	Q8C3J2
	5	363	15.2	138	11	Q9JHP2
	6	336.5	14.1	415	11	Q8JZW3
	7	334	14.0	333	11	Q9JW95
	8	331	13.9	335	11	Q8CDC3
	9	319.5	13.4	335	11	Q8CAM7
	10	317.5	13.3	333	4	Q969X2
	11	314.5	13.2	335	11	Q8CBX1
	12	308	12.9	249	4	Q9NUC5
	13	296	12.4	295	11	Q9R0G9
	14	293.5	12.3	299	4	Q9H8A2
	15	293.5	12.3	299	4	Q9ULB8
	16	163	6.8	455	4	Q96HE4
	17	163	6.8	529	4	Q8IUG7
	18	163	6.8	534	4	Q96JF0
	19	159	6.7	438	11	Q8BUU4
	20	159	6.7	451	5	Q9W121
	21	159	6.7	474	5	Q9G023
	22	143	6.0	405	6	O18974
	23	140.5	5.9	329	11	Q9WVG2
	24	140	5.9	413	4	Q8IX48
	25	140	5.9	429	4	Q8IX51
	26	140	5.9	444	4	Q8IX56
	27	139	5.8	344	4	Q8IX54
	28	139	5.8	358	11	Q9CZ48
	29	138.5	5.8	374	11	Q9DBB6
	30	137.5	5.8	277	4	Q8IX53
	31	137.5	5.8	359	4	Q8IX50
	32	137.5	5.8	359	11	Q90XF6
	33	137.5	5.8	374	11	Q922X5
	34	137.5	5.8	390	4	Q8IX58
	35	136.5	5.7	350	11	Q91WH6
	36	136	5.7	403	11	Q8K1L1
	37	135.5	5.7	349	11	Q8BSA0
	38	135.5	5.7	350	11	Q8BPL0
	39	131	5.5	224	11	Q9JL30
	40	130	5.5	325	6	Q9BEG4
	41	129.5	5.4	203	11	Q8BLV1
	42	129.5	5.4	333	11	Q91Y74
	43	129.5	5.4	333	11	Q921R5
	44	129.5	5.4	359	4	Q96L53
	45	128.5	5.4	331	11	Q8VIB3

ALIGNMENTS

RESULT 1
Q8NDV1 PRELIMINARY; PRT; 305 AA.
AC Q8NDV1
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Alpha 2,6-sialyltransferase (EC 2.4.99.7).
GN ST6GALNAC III.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Plrimates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Harduin-Leipers A.;
RT "Molecular cloning and expression of human ST6GALNAC III.";
RL Submitted (SEP-2002) to the ENBL/GenBank/DDBJ databases.
DR EMBL; AJ507291; CAD45371.1; .
DR InterPro; IPR001675; Glyco.trans.29.
DR Pfam; PF00777; Glyco.transf.29; 1.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 305 AA; 35395 MW; 700BFC6A48F8AD3A CRC64;

Alignment Scores:

Pred. No.: 5,84e-105 Length: 305
 Score: 1100.00 Matches: 207
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 46.16% Indels: 0
 DB: 4 Gaps: 0

US-09-714-936-218 (1-1294) x Q8NDV1 (1-305)
 QY 166 ATGGCGTGCATCTCTGAGGAAGTCTGTGATTCTGTGAGCTTCATAGCAGCGTTCCTT 225
 Db 1 MetAlaCysIleuLeuYsArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
 QY 226 TTCCTGCTGGTGTGGCTTGTAAATGAAGTAAATTCCTATGCTACTAACTGCTTT 285
 Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
 QY 286 GGACAACTGTGTCAGAGTGGATACCATCTCCATACATACAGAGCGGCCCTTCGAACT 345
 Db 41 GlyGlnProGlyThrIleAsnValIleProPheSerThrThrArgArgProLeuArgThr 60
 QY 346 CACTATGATACATAAATGTGAAGACACAGAGCTTCCATGCTACTGCTGCTTGT 405
 Db 61 HistyrGlyIleAsnValIleProPheSerThrThrArgArgProLeuArgThr 80
 QY 406 GCATAGTGTCAAACTCAGTGTGGTGGCCAGAGTGGGAAATGAGATAGTGA 465
 Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnValGlyAsnGluIleAsnArg 100
 QY 466 TCCTCTGCATTTGGAGATGACATGCCCCCAGAGTGGGAAATGAGATAGTGA 525
 Db 101 SerSerCysIleThrArgMetAsnAlaProThrIleValIleValIleValIleVal 120
 QY 526 CGCATGACCATGATGAGTGTGGTGGCCATACAGAGTGGGAAATGAGATAGTGA 585
 Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuAsnPro 140
 QY 586 GATTATTTTCAAGGAGCGAATACTACTATTGTGTTTATTTGGGACCTTCGCAAT 645
 Db 141 AsptyrPhePheYsGluAlaAsnThrThrIleThrValIleThrValIleThrValIle 160
 QY 646 ATGAGGAAGATGCATGCTGTCATGCTTTACACATGTTGAAAGACAGCTTGTATCTAT 705
 Db 161 MetArgYsAspGlyAsnGlyIleIleIleIleIleIleIleIleIleIleIleIleIle 180
 QY 706 CCGAATGCCAAATATACGTGACACAGAGAGCGCATGAGTACTGTGATGGAGTTT 765
 Db 181 ProAsnAlaGlnIleIleValThrThrGluYsArgMetSerThrCysAspGlyValPhe 200
 QY 766 AAGAGGAAGTGGGAGGACAGG 789
 Db 201 LysLysGluThrGlyLysAspArg 208

RESULT 2
 Q8N259 PRELIMINARY; PRT; 305 AA.
 ID Q8N259
 AC Q8N259
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ33896.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE= Tongue;
 RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,

Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 "NEDO human cDNA sequencing project.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK091215; BAC03611.1;
 DR InterPro; IPR001675; Glyco_trans_29.
 DR Pfam; PF00777; Glyco_transf_29; 1.
 KW Hypothetical protein
 SQ SEQUENCE 305 AA; 35409 MW; 7EC434C948F8B599 CRC64;
 Alignment Scores:
 Pred. No.: 7,41e-105 Length: 305
 Score: 1099.00 Matches: 206
 Percent Similarity: 99.52% Conservative: 1
 Best Local Similarity: 99.04% Mismatches: 1
 Query Match: 46.12% Indels: 0
 DB: 4 Gaps: 0

US-09-714-936-218 (1-1294) x Q8N259 (1-305)
 QY 166 ATGGCGTGCATCTCTGAGGAAGTCTGTGATTCTGTGAGCTTCATAGCAGCGTTCCTT 225
 Db 1 MetAlaCysIleuLeuYsArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
 QY 226 TTCCTGCTGGTGTGGCTTGTAAATGAAGTAAATTCCTATGCTACTAACTGCTTT 285
 Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
 QY 286 GGACAACTGTGTCAGAGTGGATACCATCTCCATACATACAGAGCGGCCCTTCGAACT 345
 Db 41 GlyGlnProGlyThrIleAsnValIleProPheSerThrThrArgArgProLeuArgThr 60
 QY 346 CACTATGATACATAAATGTGAAGACACAGAGCTTCCATGCTACTGCTGCTTGT 405
 Db 61 HistyrGlyIleAsnValIleProPheSerThrThrArgArgProLeuArgThr 80
 QY 406 GCATAGTGTCAAACTCAGTGTGGTGGCCAGAGTGGGAAATGAGATAGTGA 465
 Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnValGlyAsnGluIleAsnArg 100
 QY 466 TCCTCTGCATTTGGAGATGACATGCCCCCAGAGTGGGAAATGAGATAGTGA 525
 Db 101 SerSerCysIleThrArgMetAsnAlaProThrIleValIleValIleValIleVal 120
 QY 526 CGCATGACCATGATGAGTGTGGTGGCCATACAGAGTGGGAAATGAGATAGTGA 585
 Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuAsnPro 140
 QY 586 GATTATTTTCAAGGAGCGAATACTACTATTGTGTTTATTTGGGACCTTCGCAAT 645
 Db 141 AsptyrPhePheYsGluAlaAsnThrThrIleThrValIleThrValIleThrValIle 160
 QY 646 ATGAGGAAGATGCATGCTGTCATGCTTTACACATGTTGAAAGACAGCTTGTATCTAT 705
 Db 161 MetArgYsAspGlyAsnGlyIleIleIleIleIleIleIleIleIleIleIleIleIle 180
 QY 706 CCGAATGCCAAATATACGTGACACAGAGAGCGCATGAGTACTGTGATGGAGTTT 765
 Db 181 ProAsnAlaGlnIleIleValThrThrGluYsArgMetSerThrCysAspGlyValPhe 200
 QY 766 AAGAGGAAGTGGGAGGACAGG 789
 Db 201 LysLysGluThrGlyLysAspArg 208

RESULT 3
 Q9W6U6 PRELIMINARY; PRT; 234 AA.
 ID Q9W6U6
 AC Q9W6U6
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Alpha-N-acetylglactosamine alpha-2,6-sialyltransferase
 DE (Fragment).

GN SIAT3C.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99299247; PubMed=10369878;
RA Gilley J., Fried M.;
RT "Extensive gene order differences within regions of conserved synteny
between the Fugu and human genomes: implications for chromosomal
evolution and the cloning of disease genes.";
RL Hum. Mol. Genet. 8:1313-1320(1999).
DR EMBL; Y17466; CAB4338.1; -.
DR InterPro; IPR001675; Glyco.trans.29.
DR Pfam; PF00777; Glyco.transf.29; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER
SQ SEQUENCE 234 AA; 26937 MW; 6F9C8FCB1BC93FC8 CRC64;
Alignment Scores:
Pred. No.: 1.66e-33 Length: 234
Score: 411.00 Matches: 75
Percent Similarity: 72.79% Conservative: 24
Best Local Similarity: 55.15% Mismatches: 37
Query Match: 17.25% Indels: 0
DB: 13 Gaps: 0
US-09-714-936-218 (1-1294) x Q9W6U6 (1-234)
QY 382 TTGCACTGACGTGACCTTTGCGCATAGTGTCAAACTCAGTCAGTCAGTGGCCAG 441
Db 2 LeuAlaValHisCysAsnGlnCysAlaLeuValSerSerGlyGlnMetLeuGlyAla 21
QY 442 AAGTGGGAATGAGATAGATGATGCTCTGCTGATTTGGAGTAAGTAAGTAAGTCCCAAC 501
Db 22 GlyLeuGlyGluLeuAspLysIleGlnCysValIleArgMetAsnAlaProThr 41
QY 502 AAAGTTATGAGAAGATGTCGGCGCATGACCATGATTCGAGTTGTCGTCCTCCATPACGAG 561
Db 42 AlaGlyTyrGluGluAspValGlySerLeuThrSerLeuArgValValSerHisThrSer 61
QY 562 GTTCTCTTTGCTAAACCCCTGATTATTTTTCAGGAAGCAAGTAAGTACTATTGTT 621
Db 62 ValProLeuLeuValLysAsnGluLeuTyrTyrPheHisGlnAlaAsnThrThrTyr 81
QY 622 GTTATTTGGGACCTTTCCGCAATATGAGGAAGATGGCAATGGCATCGTTTACACAG 681
Db 82 ValPheTrpGlyProAspSerLysMetArgGlnAspGlyLysGlyGlnIlePheAsnVal 101
QY 682 TTGAAAGACAGTGTGTATCTATCCGAATGCCCAATATAGTGCACACAGAGAGCG 741
Db 102 LeuLeuLysIleAlaLysLysTyrProAsnValLysMetTyrSerMetThrSerGluLys 121
QY 742 ATGAGTACTGTGATGGATTTTAAAGGAAGCAACTGGGAAGGACAGG 789
Db 122 IleLysTyrCysAspGlnValPheGlnAsnGluThrGlyLysAsnArg 137
RESULT 4
Q8C3J2 PRELIMINARY; PRT; 302 AA.
AC Q8C3J2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sialyltransferase 7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK085730; BAC39523.1; -.
SQ SEQUENCE 302 AA; 34218 MW; 3AF7FC04F39E310B CRC64;
Alignment Scores:
Pred. No.: 2.12e-31 Length: 302
Score: 391.00 Matches: 88
Percent Similarity: 57.53% Conservative: 38
Best Local Similarity: 40.18% Mismatches: 72
Query Match: 16.41% Indels: 21
DB: 11 Gaps: 5
US-09-714-936-218 (1-1294) x Q8C3J2 (1-302)
QY 193 GTGATTGCTGTGAGCTTCATAGCAGCGTTCCTTTCTCTGCTGTGTGGTCTTGTAAAT 252
Db 14 IleLeuThrPheSerAlaValCysValPheLeu-CysCysTyrAlaCys----- 29
QY 253 GAAGTGAATTTCCCATTCCTACTAAACTGCTTTGGACACACCTGCTACAAAGTGGATACA 312
Db 30 -----LeuProLeuCysLeuAlaThrCysLeuAspArg-----H1 41
QY 313 TTCTCCT-----ACACATACAGCGCGCCCTTCGAACTCAGTATGATGATGATGATGAT 360
Db 41 SLeuProAlaAlaProArgSerThrValProGlyProLeuHisPheSer--GlyTyrSer 60
QY 361 AAGTGAAGACACAGACGCTTTCACACTGGAC----TGTGACCTTTGTGCCATGATGCA 417
Db 61 SerValProAspGlyLeuProLeuIleArgGluLeuCysHisSerCysAlaValSer 80
QY 418 AACTCAGGTTCAGATGTTGGCCAGAGTGGGAATGAGATAGATGATGATGATGATGAT 477
Db 81 SerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGlnIleAspGlyAlaGluCysVal 100
QY 478 TGGAGATGAACAATGCCCCACCAAGATTATGAGAAGATGTCGGCCGATGACCAATG 537
Db 101 LeuArgMetAsnGlnAlaProThrValGlyPheGluGluAspValGlyGlnArgSerThr 120
QY 538 ATTCTGAGTTGTCTCCATACACAGGTCCTCTTTTCTGCTAAACCCCTGATTTATTTTC 597
Db 121 LeuArgValIleSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPhe 140
QY 598 AAGGAAGCAATACTACTATTGTTGTTATTTGGGACCTTTCCGCAATATGAGGAAGAT 657
Db 141 GlnHisAlaArgAspThrLeuTyrValValTrpGlyGlnGlyArgHisMetAspArgVal 160
QY 658 GGCATGGCATCGTTTACACATGTTGAAAAGACAGTGTGTATCTATCCGATGCCAA 717
Db 161 LeuGlyGlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGln 180
QY 718 ATATACGTGACACAGAGAGCGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 181 ValTyrThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThr 200
QY 778 GGAAGACAGCGGGGCGATGCAAGCGCATCTGCTGATTTCTACACACACTTTT 828
Db 201 GlyLysAsnArgArgGlnSerGlySerPheLeu---SerThrGlyTyrPhe 216
RESULT 5
Q9JHP2 PRELIMINARY; PRT; 138 AA.
ID Q9JHP2
AC Q9JHP2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 22, Last annotation update)
DE GalNAc alpha-2,6-sialyltransferase (Fragment).


```

Db      207 PheProAsnMetGluAlaTyAlaValSerProAlaArgMetGlnGlnPheAspAspLeu 226
QY      763 TTTAAGAAGGAAGACAGAGGGCGCATGCAAGCGAGCTGCTGATTCTTACAGAC 822
      |||||: |||||:|||||:|||||: |||||: |||||: |||||: |||||:
Db      227 PheArgGlyGluThrGlyLysAspArgGlyLysSerHisSerTrpLeu---SerThrGly 245
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      823 ACTTTTATAGCGATACACAGTGCCTGGCAAGTGG-----855
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      246 TrpPheThrMetValIleAlaValGluLeuCysAspHisValHisValTyGlyMetVal 265
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      856 -----AATCACTCTTCGCGTCTTACAGC 882
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      266 ProProAspTyrcysSerGlyProAlaCysSer 276
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

RESULT 7
Q9JW95
ID      Q9JW95      PRELIMINARY;      PRT;      333 AA.
AC
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      GDI alpha/GTla alpha/GOLb alpha synthase.
GN      SIAT7F OR ST6GALNAC VI.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6; TISSUE=Brain;
RA      Okajima T., Chen H., Ito H., Kiso M., Tai T., Furukawa K., Urano T.,
RA      Furukawa K.;
RT      "Molecular cloning and expression of mouse GDI alpha/GTla alpha/GOLb
RT      alpha synthase (ST6GALNAC VI) gene.";
RL      Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB035123; BAA95940.1; -.
DR      MGD; MGI:1355316; Siat7f.
DR      InterPro; IPR001675; Glyco.trans.29.
DR      Pfam; PF00777; Glyco_transf.29; I.
SQ      SEQUENCE 333 AA; 38166 MW; D3841828D389CDEA CRC64;

Alignment Scores:
Pred. No.:      1-79e-25      Length:      333
Score:      334.00      Matches:      85
Percent Similarity:      44.68%      Conservative:      41
Best Local Similarity:      30.14%      Mismatches:      90
Query Match:      14.02%      Indels:      66
DB:      11      Gaps:      8

US-09-714-936-218 (1-1294) x Q9JW95 (1-333)
QY      73 TCTGCGGTGTACCAAGCTCCAGCTGCCCGGAGGACTGCCCTGACCCAGGCGGCC 132
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      2 AlacysSerArgProProSerGlnCysAspProThrThrLeuPro---ProGlyProPro 20
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      133 GCTGCTCGGTGGCAG-----GAGGCGCGGCGGAGCGCCATGCGCTGCATCCTGAAG 183
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      21 AlaGlyArgTrpProLeuProPheSerArgArgArgGluMetSerSerAsnLysGlu 40
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      184 AGAAAGTGTGATGTGTGAGCTTCATACAGCGTTCCTTTTCCTGCTGCTGTGCGGT 243
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      41 GlnArgSerAlaValPheVal-----IleLeuPheAlaLeuIleThrIle 55
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      244 CTTGTGTA-----AATGAAGTGAATTCCTTCCATGCTACTAAATGC 282
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      56 LeuIleLeuTySerSerAsnSerAlaAsnGluVal-----67
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      283 TTTGGACAACCTGTGTACAAAGTGGATACCATCTCTTACATACAGGCGGCCCTTCGA 342
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      67 -----67
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      343 ACTCACTATGATACATAAATGTGAAGACACAGAGCCTTTGCACTGCAC-----393
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

```

```

Db      68 PheHisTyrglySerLeuArgGlyArgThrArgArgProValAsnLeuLysLysTrpSer 87
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      394 -----TGTGACCTT 402
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      88 PheSerSerAlaTyrcysSerGlyProAlaCysSer 107
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      403 TGTGCCCATAGTGTCAAACTCAGGTTCAGATGGTGGCCAGAGGTGGGAATAGATAGAT 462
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      108 CysValIleIleThrSerSerHisLeuLeuGlyThrLysLeuGlyProGluIleGlu 127
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      463 CGATCTCTCTGCTGATTTGGAGAATGAACAATGCCCCACCAAGGTTATGAAGAAGATGTC 522
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      128 ArgAlaGluCysThrIleArgMetAsnAspAlaProThrSerGlyTyrcysSerAlaAspVal 147
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      523 GGCCGCATGACCATGATTCGAGTTGTGTCCCATACCAGCGTTCCTTTTGTCTAAATAAC 582
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      148 GlyAsnLysThrThrPheArgValAlaHisSerSerValPheArgValLeuArgLys 167
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      583 CCTGATTTATTTTTCAGGAAGCGAATACTACTATTGTGTATTTCGGGACCTTCCGC 642
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      168 ProGlnGluPheValAsnArgThrProGluThrValPheIlePheTrpGlyProProAsn 187
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      643 AATATGAGGAAAGATGGCAATGGCATCGTTTACAAACATGTTGAAAAAGACAGTGTGATC 702
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      188 LysMetGlnLysPro---GlnGlySerLeuLeuArgValIleGlnArgAlaGlyLeuMet 206
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      703 TATCCGAATGCCCAATATACGTGACACAGAGAGCGCATGAGTTACTGTGTGAGGAGTT 762
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      207 PheProAsnMetGluAlaTyAlaValSerProAlaArgMetGlnGlnPheAspAspLeu 226
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      763 TTTAAGAAGGAAGTGGGAAGCAGAGGGCGCATGCAAGGCGACTGCTGATTTCTACAGAC 822
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      227 PheArgGlyGluThrGlyLysAspArgGluLysSerHisSerTrpLeu---SerThrGly 245
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY      823 ACTTTT 828
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      246 TrpPhe 247
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

RESULT 8
Q8CDC3
ID      Q8CDC3      PRELIMINARY;      PRT;      335 AA.
AC
DT      01-WAR-2003 (TrEMBLrel. 23, Created)
DT      01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Sialyltransferase 7.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
DR      EMBL; AK030648; BAC27064.1; -.
SQ      SEQUENCE 335 AA; 38477 MW; B0D1B9E878501D78 CRC64;

Alignment Scores:
Pred. No.:      3-67e-25      Length:      335
Score:      331.00      Matches:      83
Percent Similarity:      45.22%      Conservative:      40
Best Local Similarity:      30.51%      Mismatches:      83
Query Match:      13.89%      Indels:      66
DB:      11      Gaps:      8

US-09-714-936-218 (1-1294) x Q8CDC3 (1-335)
QY      103 CCCAGGACTGCCCTGACCCAGGCGCGCGTGTCTGTGTCAG-----GAGGGC 153

```

```

Db      14  ProThrThrLeuPro---ProGlyProProIaGlyArgTrpProLeuProPheSerArg 32
QY      154  CGCGGAGCGCCATGCGCTGCAFCCTGAAGAAAGTCTGTGATTCTGTGAGCTTCATA 213
Db      33  ArgArgArgGluMetSerSerAsnLysGluGlnArgSerAlaValPheVal-----49
QY      214  GCAGGGTTCCTTTCCCTGCTGGTGGCTGCTTCTA-----AAT 252
Db      50  -----IleLeuPheAlaIleuIleThrIleLeuIleLeuTyrSerSerAsnSerAlaAsn 67
QY      253  GAAGTGAATTTCCCATCTTACTAAACTGCTTTGGCAACCTGCTGCAACCTGCTGCAACCTGATACCA 312
Db      68  GluVal-----69
QY      313  TTCTCTACATACACAGCGGCGCCCTTCCGACTCCTATGATGATACATAAATGTGAAGACA 372
Db      70  -----PheHisTyrGlySerLeuArgValArgThr 79
QY      373  CAAGAGCCTTTGCAACTGGAC-----393
Db      80  ArgArgProValAsnLeuLysLysTrpSerPheSerSerAlaTyrPheProIleuGly 99
QY      394  -----TGTGACCTTTGTGCCATAGTGTCAAACTCAGGTCAAGATG 432
Db      100  AsnLysThrLeuProSerArgCysAsnGlnCysValIleIleThrSerSerHisLeu 119
QY      433  GTTGCCAGAGGTGGGAATGATAGATGATCGATCCCTCCGATTTGGAGATGAACAAT 492
Db      120  LeuGlyThrLysLeuGlyProGluIleGluArgAlaGluCysThrIleArgMetAsnAsp 139
QY      493  GCCCCACCAAGGTTATGAAGATGTCCGCGCGCATGACCATGATTCGAGTTCTGTCC 552
Db      140  AlaProThrSerGlyTyrSerAlaAspValGlyAsnLysThrThrPheArgValAla 159
QY      553  CATACACGTTCTCTTTGCTTAAACCCCTGATTTATTTTCAAGGAGCGAATACT 612
Db      160  HisSerSerValPheArgValLeuArgLysProGlnPheValAsnArgThrProGlu 179
QY      613  ACTATTGTGTTATTTGGGACCTTTCCGCAATATGAGGAAGATGCCAATGGCATGTT 672
Db      180  ThrValPheIlePheTrpGlyProProAsnLysMetGlnLysPro---GlnGlySerLeu 198
QY      673  TACAACATGTTGAAAAAGACAGTGTGTATCTATCCGAATCCCAATATACATGACACACA 732
Db      199  LeuArgValIleGlnArgAlaGlyLeuMetPheProAsnMetGluAlaTyrAlaValSer 218
QY      733  GAGAAGCGCAGTACTTACTGTGATGAGCTTTTAAAGAGGAAGAACTGGGAGGACAGGGG 792
Db      219  ProAlaArgMetGlnGlnPheAspLeuPheArgGlyGluThrGlyLysAspArgGlu 238
QY      793  CATCAAGCGCAGTCTGCTGATTTCTACAGACACTTTT 828
Db      239  LysSerHisSerTrpLeu---SerThrGlyTrpPhe 249

```

RESULT 9

```

Q8CAM7  PRELIMINARY; PRT; 335 AA.
AC Q8CAM7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Hypothetical protein (Similar to sialyltransferase 7 ((alpha-N-
DE acetylneuraminyl 2,3-betagalactosyl-1,3)-N-acetyl galactosaminide
DE alpha-2,6-sialyltransferase) F).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA THE RIKEN Genome Exploration Research Group Phase I & II Team;

```

```

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK038434; BAC29997.1; -.
SQ SEQUENCE 335 AA; 38333 MW; E3295C30BAD77B07 CRC64;

Alignment Scores:
Pred. No.: 5,73e-24 Length: 335
Score: 319.50 Matches: 68
Percent Similarity: 58.52% Conservative: 35
Best Local Similarity: 38.64% Mismatch: 64
Query Match: 13.41% Indels: 9
DB: 11 Gaps: 2

US-09-714-936-218 (1-1294) x Q8CAM7 (1-335)
QY 278 ACTGCTTTGGACAACCTGTCACAAAGTGGATPACCATTCCTCT-----ACACAT 325
Db 53 ThrGlySerThrGlnLeuValGluSer-----SerProGlnProArgArgThrAla 69
QY 326 ACAGCGCGCCCTTCGAACCTCCTATGATGATACATAAATGTGAAGACACAGAGCTTTGC 385
Db 70 ProAlaGlyProArgGlnLeuGlu--GlyTyrLeuGlyValAlaAspHisLysProLeuL 89
QY 386 AACTGGACTGTGACCTTTGTCCATAGTGTCAAACTCAGGTGATGTTGGCCAGAGG 445
Db 89 ysmetHisCysLysAspCysAlaLeuValThrSerSerGlyHisLeuLeuArgSerGln 109
QY 446 TGGGAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 505
Db 109 InGlyProHisIleAspGlnThrGluCysValIleArgMetAsnAspAlaProThrArg 129
QY 506 GTTATGAGAAGATGTCGCGCCATGACCATGATGATGATGATGATGATGATGATGATG 565
Db 129 IyTyrGlyLeuAspValGlyAsnArgThrSerLeuArgValIleAlaHisSerIleG 149
QY 566 CTCTTTTCTTAAAAACCTGATTTATTTTCAAGAGACGCAATACTACTATTTGTGTTA 625
Db 149 InaGlyIleLeuArgAsnArgHisAspLeuLeuAsnValSerGlnGlyThrValPheIle 169
QY 626 TTTGGGACCTTTCCGCAATATGAGGAAGATGCAATGGCATGCGTTCATACATGTTGA 685
Db 169 heTrpGlyProSerSerTyrMetArgAspGlyLysGlyGlnAlaTyrAsnAsnLeuG 189
QY 686 AAAGACAGTGTGATCTATCCGAATGCCAAATATACGTGACCAAGAGAGCGCATGA 745
Db 189 InLeuLeuSerGlnValLeuProArgLeuLysAlaPheMetIleThrArgHisArgMetL 209
QY 746 GTTACTGTGTGAGTGTGTTTAAAGAGAACTGGGAGGACAGG 789
Db 209 euGlnPheAspGluPheLeuGlnGluThrGlyLysAspArg 223

RESULT 10
Q969X2  PRELIMINARY; PRT; 333 AA.
ID Q969X2
AC Q969X2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Hypothetical protein (Similar to sialyltransferase 7 ((alpha-N-
DE acetylneuraminyl 2,3-betagalactosyl-1,3)-N-acetyl galactosaminide
DE alpha-2,6-sialyltransferase) F).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Muscle;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016299; AAH16299.1; -.
DR EMBL; BC006564; AAH06564.1; -.

```



```

[1]
RN  SEQUENCE FROM N.A.
RP  TISSUE=Thyroid;
RA  Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA  Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA  Tanase T., Nomura Y., Tojiya S., Komai F., Hara R., Takeuchi K.,
RA  Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA  Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;
RT  "NEDO human cDNA sequencing project.";
RL  Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK023900; BAB14715.1; -.
DR  InterPro; IPR001675; Glyco.trans.29.
DR  Pfam; PF00777; Glyco.transf_29; 1.
KW  Hypothetical protein
SQ  SEQUENCE 299 AA; 34261 MW; 10C27604122F4BBD CRC64;

Alignment Scores:
Pred. No.:      2,77e-21      Length:      299
Score:          293.50      Matches:      74
Percent Similarity: 47.58%      Conservative: 34
Best Local Similarity: 32.60%      Mismatches:  88
Query Match:    12.32%      Indels:      31
DB:              4          Gaps:         6

US-09-714-936-218 (1-1294) x Q9H8A2 (1-299)
QY  181 AAGAGAAAGTCGTGATTCGTGAGCTTC---ATAGCAGCGTTCCTTTTCTCTGCTGGTT 237
Db  7 GlnArgSerAlaValPheValIleLeuPheAlaLeuIleThrIleLeuIleLeuTyrSer 26
QY  238 GTGCGTCTTGTAATGAAGTGAATTTCCCATTCCTACTAACTGCTTTGGACACCTGGT 297
Db  27 SerAsnSerAlaAsnGluVal----- 33
QY  298 ACAAGTGGATACCATTCCTCTAC-----ACATACAGCGCGCCCTTCGA 342
Db  34 -----PheHisTyrGlySerLeuArgGlyArgSerArgArgProValAsn 48
QY  343 ACTCACTAT-----GGATACATAAATGTCAGACACACAGACCCCTTGCAG 387
Db  49 LeuLysLysTrpSerIleThrAspGlyTyrValProIleLeuGlyAsnLysThrLeuPro 68
QY  388 CTGGACTGTCACCTTTGGCCATAGTCACAACTCAGGTCCAGATGTTGGCCAGAGGTG 447
Db  69 SerArgCysHisGlnCysValIleValSerSerSerHisLeuLeuGlyThrLysLeu 88
QY  448 GGAATGAGATAGATCGATTCCTCTGCGATTGGAGAAATGAACAATGCCCCACCAAGGT 507
Db  89 GlyProGluIleGluArgAlaGluCysThrIleArgMetAsnAspAlaProThrThrGly 108
QY  508 TATCAAGAGATGTCGCGCGCATGACCATGATTCGAGTTGTCCTCCATACACGAGTTCCT 567
Db  109 TyrSerAlaAspValGlyAsnLysThrThrTyrArgValValAlaHisSerValPhe 128
QY  568 CTTTGTGTAATAAACCCCTGATTATTTCACGAAAGCAGATCTACTATTGTGTATT 627
Db  129 ArgValLeuArgArgProGlnGluPheValAsnArgThrProGluThrValPheIlePhe 148
QY  628 TGGGAGCTTCCGCAATATGAGGAAGATGGCAATGGCATCGTTTACACATCTGAAA 687
Db  149 TrpGlyProSerLysMetGlnLysPro---GlnGlySerLeuValArgValIleGln 167
QY  688 AAGACAGTGTGATCTATCCGATGCCCAATATAGTGTACACACAGAGACGATAGT 747
Db  -168 ArgAlaGlyLeuValPheProAsnMetGluAlaTyrAlaValSerProGlyArgMetArg 187
QY  748 TACTGTGATGGATCTTTTAAAGAGGAACTGGGAAGCAGAGGGCGCATCGAAGCGACTG 807
Db  188 GlnPheAspAspLeuPheArgGlyGluThrGlyLysAspArgGlnLysSerHisSerTrp 207
QY  808 CTGATTCTACAGACACTTTT 828
Db  208 Leu---SerThrGlyTrpPhe 213

```

```

RESULT 15
Q9ULB8
ID  Q9ULB8      PRELIMINARY;      PRT;      299 AA.
AC  Q9ULB8;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  N-acetylgalactosaminide alpha2,6-sialyltransferase.
GS  ST6GALNAC VI.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Yoshida A.;
RT  "human N-acetylgalactosaminide alpha2,6-sialyltransferase.";
RL  Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Livers A.;
RT  "Molecular cloning and expression of human ST6GALNAC VI.";
RL  Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB035173; BAB87035.1; -.
DR  EMBL; AJ507293; CAD45373.1; -.
DR  InterPro; IPR001675; Glyco.trans.29.
DR  Pfam; PF00777; Glyco.transf_29; 1.
KW  Glycosyltransferase; Transferase.
SQ  SEQUENCE 299 AA; 34291 MW; F9DAFF04008A0C5F CRC64;

Alignment Scores:
Pred. No.:      2,77e-21      Length:      299
Score:          293.50      Matches:      74
Percent Similarity: 47.58%      Conservative: 34
Best Local Similarity: 32.60%      Mismatches:  88
Query Match:    12.32%      Indels:      31
DB:              4          Gaps:         6

US-09-714-936-218 (1-1294) x Q9ULB8 (1-299)
QY  181 AAGAGAAAGTCGTGATTCGTGAGCTTC---ATAGCAGCGTTCCTTTTCTCTGCTGGTT 237
Db  7 GlnArgSerAlaValPheValIleLeuPheAlaLeuIleThrIleLeuIleLeuTyrSer 26
QY  238 GTGCGTCTTGTAATGAAGTGAATTTCCCATTCCTACTAACTGCTTTGGACACCTGGT 297
Db  27 SerAsnSerAlaAsnGluVal----- 33
QY  298 ACAAGTGGATACCATTCCTCTAC-----ACATACAGCGCGCCCTTCGA 342
Db  34 -----PheHisTyrGlySerLeuArgGlyArgSerArgArgProValAsn 48
QY  343 ACTCACTAT-----GGATACATAAATGTCAGACACACAGACCCCTTGCAG 387
Db  49 LeuLysLysTrpSerIleThrAspGlyTyrValProIleLeuGlyAsnLysThrLeuPro 68
QY  388 CTGGACTGTCACCTTTGGCCATAGTCACAACTCAGGTCCAGATGTTGGCCAGAGGTG 447
Db  69 SerArgCysHisGlnCysValIleValSerSerSerHisLeuLeuGlyThrLysLeu 88
QY  448 GGAATGAGATAGATCGATTCCTCTGCGATTGGAGAAATGAACAATGCCCCACCAAGGT 507
Db  89 GlyProGluIleGluArgAlaGluCysThrIleArgMetAsnAspAlaProThrThrGly 108
QY  508 TATCAAGAGATGTCGCGCGCATGACCATGATTCGAGTTGTCCTCCATACACGAGTTCCT 567
Db  109 TyrSerAlaAspValGlyAsnLysThrThrTyrArgValValAlaHisSerValPhe 128
QY  568 CTTTGTGTAATAAACCCCTGATTATTTCACGAAAGCAGATCTACTATTGTGTATT 627
Db  129 ArgValLeuArgArgProGlnGluPheValAsnArgThrProGluThrValPheIlePhe 148

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:29:54 ; Search time 36 Seconds

(without alignments)

6913.460 Million cell updates/sec

Title:

US-09-714-936-218

Perfect score: 2383

Sequence: 1 ccggaattccgggtcgagc.....ttttctctctttttttt 1294

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched:

283308 seqs, 9618682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09714936/runat_10092003_151459_5705/app_query.fasta_1.1479
-DB=pir76 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -SPAR=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=prt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORF=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714936 -CGN_1_1_35=brunat_10092003_151459_5705 -NCP=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSEBLOCK=100 -IONGLOS
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

pir76:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	163	6.8	413	2 S41114	Gal beta 1,4 GlnNA
2	142	6.0	406	2 A41734	beta-galactoside a
3	138.5	5.8	374	2 A45074	Gal beta 1,3(4)Glc
4	138.5	5.8	403	2 A28451	beta-galactoside a
5	137.5	5.8	375	2 JN0618	Gal beta 1,3(4)Glc
6	136	5.7	342	2 S55675	Gal-beta-1,3GalNAc
7	135.5	5.7	350	2 B54420	beta-galactoside a
8	135.5	5.7	350	2 A54420	beta-galactoside a
9	134	5.6	340	2 I54229	beta-galactoside a
10	131.5	5.5	350	2 JC5251	beta-galactoside a
11	131	5.5	337	2 S36824	beta-galactoside a
12	126	5.3	380	2 JC6321	CMP-NeuAc-GM3 alph
13	125.5	5.3	343	2 A45073	Gal beta 1,3GalNAc
14	124	5.2	380	2 A56950	alpha-2,8-sialyltra

15	123	5.2	398	2 G86216	protein T23G18.14
16	120	5.0	470	2 T49278	hypothetical prote
17	118.5	5.0	329	2 A48715	gal-beta(1-3/1-4)G
18	118.5	5.0	332	2 A49879	alpha-2,3-sialyltr
19	116.5	4.9	404	2 A54871	Gal beta-1, 3GalNA
20	113	4.7	480	2 JC7812	BGL6 homologous xi
21	112	4.7	526	2 JC7248	alpha-N-acetylglal
22	111	4.7	359	2 JC4973	GM3 alpha-2,8-sial
23	110	4.6	387	2 JC0364	lactosylceramide a
24	110	4.6	566	2 A49880	alpha-N-acetylglal
25	108	4.5	342	2 JC5195	alpha-N-acetylneur
26	106	4.4	356	2 A54032	alpha-N-acetylneur
27	103	4.3	375	2 I48686	N-glycan alpha 2,8
28	101	4.2	375	2 I39169	sialyltransferase
29	100	4.2	222	2 B54898	STX protein - huma
30	100	4.2	375	2 A46727	sialyltransferase
31	97.5	4.1	425	2 T00720	hypothetical prote
32	95	4.0	830	2 T36683	probable integral
33	92.5	3.9	1023	2 T30257	IgG Fc binding pro
34	92	3.9	359	2 I39403	alpha-2,8-polysial
35	92	3.9	1053	2 T51016	serine proteinase
36	92	3.9	1548	2 S34583	probable gpase-ac
37	91	3.9	1822	2 T14106	hypothetical prote
38	89.5	3.8	418	2 T27506	hypothetical prote
39	89	3.7	633	2 T47524	primase - ateline
40	88	3.7	835	2 T42970	gamma-crystallin 4
41	87.5	3.7	174	1 CYMSG4	natural killer cel
42	87	3.7	216	2 PT0375	alpha-N-acetylneur
43	87	3.7	359	2 JC4224	polysialyltransfer
44	87	3.7	359	2 S52425	ubiquinol-cytochro
45	86.5	3.6	257	2 T51305	

ALIGNMENTS

RESULT 1

S41114

Gal beta 1,4 GlnNAc alpha 2,6-sialyltransferase - chicken

C:Species: Gallus gallus (chicken)

C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: S41114

R:Kurosawa, N.; Kawasaki, M.; Hamamoto, T.; Nakaoaka, T.; Lee, Y.C.; Arita, M.; Tsuj:

Eur. J. Biochem. 219, 375-381, 1994

A:Title: Molecular cloning and expression of chick embryo Gal-beta-1,4GlcNAc-alpha-

A:Reference number: S41114; PMID:8307003

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-413 <KUR>

A:Cross-references: EMBL:X75558; NID:9452750; PIDN:CAA53235.1; PID:9452751

C:Superfamily: beta-galactoside alpha-2,6-sialyltransferase

Alignment Scores:

Pred. No.:	1.93e-06	Length:	413
Score:	163.00	Matches:	47
Percent Similarity:	40.68%	Conservative:	25
Best Local Similarity:	26.55%	Mismatches:	47
Query Match:	6.84%	Indels:	59
DB:	2	Gaps:	4

US-09-714-936-218 (1-1294) x S41114 (1-413)

Qy	343	ACTCACTATGGATACATAAATGTGAAGACACAAGACGCTTTGCAACTGACTGTGACCTT	402
		: :	
Db	131	fhrryrasnGlyysMetasnaalalalysleuSerProgluGlnleuLeuCysArgleu	150
		: :	
Qy	402	-----	402
Db	151	ArgaspargValasnValThrMetileArgglySerAspGlyProPheasnSerSerGlu	170
Qy	403	-----	403
Db	171	TrpGlnHisTryleuProasplysSerleuAsnGluThrValglyArgleuGlyArgcys	190
		-----TGT	405

[illegible]

QY 343 ACTCACTATGGATACATAAATGTTGAAGACACAAAGACGCTTTG----- 384
 Db 121 GluPheValProPheGlyIleLysGlyClnAspAsnLeuIleLysAlaIleLeuSer 140
 QY 385 -----CAACTGGACTGTGACCTTTGTGCC 408
 Db 141 ValThrLysGluTyrArgLeuThrProAlaLeuAspSerLeuHisCysArgGlyCysIle 160
 QY 409 ATAGTGTCAAACTCAGGTCCAGATGGTTGGCCAGCAAGGTGGAAATGAGATAGATCGATCC 468
 Db 161 IleValGlyAsnGlyValLeuAlaAsnLysSerLeuGlySerArgIleAspAspTyr 180
 QY 469 TCCTGCATTGGAGATGAACAATGCCGCCACCAAAAGGTTATGAAGAAGATGTGGCGCG 528
 Db 181 AspileValIleArgLeuAsnSerAlaProValLysGlyPheGlyLysAspValGlySer 200
 QY 529 ATGACCATGATTGAGTGTGTGTGCCATACCAAGGTTCCCTCTTTTCTAAAAAACCTGAT 588
 Db 201 LysThrThrLeuArgIle-----ThrTyrProGluGlyAlaMetGlnArgProGlu 217
 QY 589 TATTTTTCAGGAACGGAATACTACTATTGTGTTATT---TGGGGACCTTCCGCAAT 645
 Db 218 GlnTyrGluArgAspSerLeuPheValLeuAlaGlyPheLysTyrGlnAspPheLysTyr 237
 QY 646 ATGAGGAAGATGGCAATGGCATCGCTTTACAAATGTTGAAAAGACAGATTGGTATCTAT 705
 Db 238 LeuLys----- 239
 QY 706 CCGAATGCCAAATATACGTACCCACAGACAGAGCGGATGATGATGTGATGGAGTTT 765
 Db 240 -----TyrIleValTyrLysGluArgValSerAlaSerAspGlyPheTyr 254
 QY 766 AAG 768
 Db 255 Lys 255

RESULT 4
 A28451
 beta-galactoside alpha-2,6-sialyltransferase (EC 2.4.99.1) - rat
 N;Alternate names: CMP-N-acetylneuraminase-alpha-N-acetyl-galactosamide-alpha-2,6-sialy
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 24-Sep-1999
 R;Weinstein, J.; Lee, E.U.; McEntee, K.; Lai, P.H.; Paulson, J.C.
 J. Biol. Chem. 262, 17735-17743, 1987
 A;Title: Primary structure of beta-galactoside alpha-2,6-sialyltransferase. Conversion c
 A;Reference number: A28451; PMID:86087067; PMID:3121604
 A;Accession: A28451
 A;Molecule type: mRNA
 A;Residues: 1-403 <WE1>
 A;Cross-references: GB:M18769; NID:g204254; PIDN:AAA41196.1; PID:g204255
 A;Note: the authors translated the codon GAC for residue 367 as His
 R/O'Hanlon, T.P.; Lau, K.M.; Wang, X.; Lau, J.T.Y.
 J. Biol. Chem. 264, 17389-17394, 1989
 A;Title: Tissue-specific expression of beta-galactoside alpha-2,6-sialyltransferase. Tra
 A;Reference number: A34465; PMID:90008905; PMID:2793863
 A;Accession: A34465
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 183-245,252-282 <OHA>
 A;Experimental source: renal
 C;Superfamily: beta-galactoside alpha-2,6-sialyltransferase
 C;Keywords: glycosyltransferase; transmembrane protein

Alignment Scores:
 Pred. No.: 0.000328 Length: 403
 Score: 138.50 Matches: 38
 Percent Similarity: 45.19% Conservative: 23
 Best Local Similarity: 28.15% Mismatches: 47
 Query Match: 5.81% Indels: 27
 DB: 2 Gaps: 6

US-09-714-936-218 (1-1294) x A28451 (1-403)

QY 238 GTGCGTCTTGAATGAAGTGAATTTCCATTCCTAACTGCTTTGGACACCTGGT 297
 Db 147 ValSerMetIleGluAlaThrAspPhePhe-----AsnThr 159
 QY 298 ACAAAGTGGATACCAATTCCTACACATACAGCGGCCCTTCGAACACTCACTATGATAC 357
 Db 160 ThrGluTyr-----GluGlyTyrLeuProLysGluAsnPheArgThrLysValGly--- 176
 QY 358 ATAAATGTGAAGACAAAGAGCCCTTCAACTGGACTGTGACCTTTGTGCCATAGTGTCA 417
 Db 177 -----ProTyrPheLys-----ArgCysAlaValValSer 185
 QY 418 AACTCAGCTCAGATGTTGGCCAGAGAGTGGGAATGAGATAGATCGATCTCTCGATT 477
 Db 186 SerAlaGlySerLeuLysAsnSerGlnLeuGlyArgGluIleAspAsnHisAspAlaVal 205
 QY 478 TGGAGAATGAACAATGCCGCCACCAAAAGGTTATGAAGAAGATGTGGCGCGCATGACCATG 537
 Db 206 LeuArgPheAsnGlyAlaProThrAspAsnPheGlnGlnAspValGlySerLysThrThr 225
 QY 538 ATTGAGTGTGTGCCATACAGCGTTCCTCTTTGCTTAAAAACCCGATATATTTTC 597
 Db 226 IleArgLeuMetAsnSerGln-----LeuValThrThrGluLysArgPheLeu 241
 QY 598 AAGGACGCGAATACTACT-----ATTGTGTTATTGGGACCT 636
 Db 242 LysAspSerLeuTyrThrGluGlyIleLeuIleValTyrAspPro 256

RESULT 5
 JN0618
 Gal beta 1,3(4)GlcNAc alpha-2,3-sialyltransferase (EC 2.4.99.-) - human
 C;Species: Homo sapiens (man)
 C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 28-May-1999
 C;Accession: JN0618
 R;Kitagawa, H.; Paulson, J.C.
 Biochem. Biophys. Res. Commun. 194, 375-382, 1993
 A;Title: Cloning and expression of human Galbeta1.3(4)GlcNAc alpha2,3-sialyltransfe.
 A;Reference number: JN0618; PMID:9326146; PMID:8333853
 A;Accession: JN0618
 A;Molecule type: mRNA
 A;Residues: 1-375 <KIT>
 A;Cross-references: GB:L23768; NID:g388014; PIDN:AAA35778.1; PID:g388015
 C;Comment: This enzyme catalyzes the transfer of sialic acid to the terminal positi
 C;Superfamily: alpha-2,3-sialyltransferase 57%
 C;Keywords: glycosyltransferase

Alignment Scores:
 Pred. No.: 0.000403 Length: 375
 Score: 137.50 Matches: 39
 Percent Similarity: 40.33% Conservative: 34
 Best Local Similarity: 21.55% Mismatches: 63
 Query Match: 5.77% Indels: 45
 DB: 2 Gaps: 4

US-09-714-936-218 (1-1294) x JN0618 (1-375)

QY 283 TTTGGACAACCTGGTACAAAGTGGATACCATCTCTCTACACATACAGCGGCCCTTCGA 342
 Db 102 PheSerLysProAlaProMetPheLeuAspSerPheArgLysTyrAlaArgIleArg 121
 QY 343 ACTCAGTATGGATACATAAATGTAAGACAAAGACGCTTTG----- 384
 Db 122 GluPheValProPheGlyIleLysGlyGlnAspAsnLeuIleLysAlaIleLeuSer 141
 QY 385 -----CAACTGGACTGTGACCTTTGTGCC 408
 Db 142 ValThrLysGluTyrArgLeuThrProAlaLeuAspSerLeuArgCysArgArgCysIle 161
 QY 409 ATAGTGTCAAACTCAGGTCCAGATGGTTGGCCAGCAAGGTGGAAATGAGATAGATCGATCC 468
 Db 162 IleValGlyAsnGlyValLeuAlaAsnLysSerLeuGlySerArgIleAspAspTyr 181

```
QY 469 TCTGCGATTGGAGAAATGAACAATGCCCCACCAAGGTTATGAGAAGATGCGCGCGC 528
   :: |||::|||::|||:: |||::|||::|||:: |||::|||::|||:: |||::|||::
Db 182 AspIleValValArgLeuAsnSerAlaProValIysGlyPheGluIysAspValGlySer 201
QY 529 ATGACCATGATTCGAGTGTGCTCCATACACCGCTCTCTCTTTGTCTAAACACCTCAT 588
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 LysThrThrLeuArgIle-----ThrTyrProGluGlyAlaMetGlnArgProGlu 218
QY 589 TATTTTTCACGAAGCGAATACTACTATTGTGTATT--TGGGACCTTCCGCAAT 645
   :: |||::|||::|||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 GlnTyrGluArgSerLeuPheValLeuAlaGlyPheLysTrpGlnAspPheLysTrp 238
QY 646 ATGAGGAACAGATCGCAATGCGATCGTTTACACATGTTGAAAAAGACAGTGTGATCTAT 705
   ::|||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 LeuLys-----240
QY 706 CCGAATGCCAAATATAGTCACACAGAGAGCGATGAGTACTGTGATGAGCTTTT 765
   |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Db 241 -----TyrIleValTyrLysGluArgValSerAlaSerAspGlyPheTrp 255
QY 766 AAG 768
   |||
Db 256 Lys 256

RESULT 6
S55675
Gal-beta-1,3GalNac alpha-2,3-sialyltransferase - chicken
C:Species: Gallus gallus (chicken)
C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Sep-1999
C:Accession: S55675
R:Kurosawa, N.; Hamamoto, T.; Inoue, M.; Tsuji, S.
Biochim. Biophys. Acta 1244, 216-222, 1995
A:Title: Molecular cloning and expression of chick Gal-beta-1,3GalNac alpha-2,3-sialyltr
A:Reference number: S55675; MUID:95284088; PMID:7766661
A:Accession: S55675
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-342 <KUR>
A:Cross-references: GB:X80503; NID:g975654; PIDN:CAA56666.1; PID:g975655
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase

Alignment Scores:
Pred. No.: 0.00055 Length: 342
Score: 136.00 Matches: 42
Percent Similarity: 44.36% Conservative: 17
Best Local Similarity: 31.58% Mismatches: 54
Query Match: 5.71% Indels: 20
DB: 2 Gaps: 4

US-09-714-936-218 (1-1294) x S55675 (1-342)
QY 373 CAAGAGCCTTTGCAA-----CTGGACTGTGACCTTTGTGGCATAGTGTCAAAC 420
   ::|||::|||::|||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 ArgAspProLeuGlnGluArgGlyThrPheSerCysArgGlyCysAlaValValGlyAsn 149
QY 421 TCAGTCAGATGTTGGCCAGAGGTGGAAATGAGATAGATGATGATCTCTCATTTGG 480
   ||| ||| :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 SerGlyAsnLeuArgGlnSerGlnTyrGlnAspIleAspSerHisAspPheValLeu 169
QY 481 AGAATGACAAATGCCGCCACCAAGGTTATGAGAAGATGTGGCGCGCATGACCATGATT 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 ArgMetAsnArgAlaProThrIleGlyTyrGluSerAspValGlySerLysThrHis 189
QY 541 CGAGTTGTGTCATACACAGGCTCTCTTTGCTGCTAAAAACCTGATTTATTTTCAAG 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 HisPheValTyrProGluSer-----TyrLys 198
QY 601 GAAGCGAATACTACTATTTCGTATTATTTGGGACCTTCCGCAATATGAGAAGATGCG 560
   ||| ||| :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 GluLeuAlaGluAsnValSerMetIleValIleProPheLysThrLeu-----AspLeu 216
QY 661 AATGCGATCGTTTACACATGTTGAAAAAGACAGTGGT-----ATCATCCGAAT 711
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 217 ArgTrpIleValThrAlaLeuThrThrGlyThrIleAsnPheThrTyrValProValPro 336
QY 712 GCCCAATATACGTGACCCACAGAGAGCGCATGAGTTAC 750
   ::|||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 ArgLysIleLysValArgLysGluLysValLeuIleTyr 249

RESULT 7
B54420
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3Gala.2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Sep-1999
C:Accession: B54420
R:Lee, Y.C.; Kojima, N.; Wada, E.; Kurosawa, N.; Nakaoka, T.; Tsuji, S.
J. Biol. Chem. 269, 10028-10033, 1994
A:Title: Cloning and expression of cDNA for a new type of Galbeta1,6GalNAc alpha2,3-
A:Reference number: A54420; MUID:94193584; PMID:8144500
A:Accession: B54420
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-350 <LEE>
A:Cross-references: GB:X76988; NID:g475225; PIDN:CAA54293.1; PID:g475226
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
C:Keywords: glycosyltransferase; transmembrane protein

Alignment Scores:
Pred. No.: 0.000611 Length: 350
Score: 135.50 Matches: 34
Percent Similarity: 51.11% Conservative: 12
Best Local Similarity: 37.78% Mismatches: 31
Query Match: 5.69% Indels: 13
DB: 2 Gaps: 4

US-09-714-936-218 (1-1294) x B54420 (1-350)
QY 373 CAAGAGCCTTTGCAACTGCGACTGTGACCTTTGTGCCATAGTGTCAACTCAGGTCAATG 432
   ::|||::|||::|||:: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 ArgAspProGlnGln-----CysArgArgCysAlaValValGlyAsnSerGlyAsnLeu 161
QY 433 GTTGCCAGAAAGGTGGAAATGAGATGATGATGATCTCTCTGCTTTGGAGATGAACAAT 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 ArgGlySerGlyTyrGlyGlnGluValAspSerHisAsnPheIleMetArgMetAsnGln 181
QY 493 GCCCCACCAAGGTTATGAGAAGATGTCGCGCGCATGATGATGATGATGATGATGATGATG 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 AlaProThrValGlyPheGluLysAspValGlySer-----ArgThrThrHis 197
QY 553 CATACACGCTTCCTCTTTGCTAAACACCTGATTTATTTTCAAGGACGAAATACT 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 HisPheMetTyrProGluSerAlaLysAsn-----LeuProIleAsnVal 212
QY 613 ACTATTGTGTTATTTGGGACCTTTCCGC 642
   :: |||::|||::|||:: ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 SerPheValLeuVal-----ProPheLys 220

RESULT 8
A54420
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3Gala.2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Sep-1999
C:Accession: A54420
R:Lee, Y.C.; Kojima, N.; Wada, E.; Kurosawa, N.; Nakaoka, T.; Tsuji, S.
J. Biol. Chem. 269, 10028-10033, 1994
A:Title: Cloning and expression of cDNA for a new type of Galbeta1,6GalNAc alpha2,3-
A:Reference number: A54420; MUID:94193584; PMID:8144500
A:Accession: A54420
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-350 <LEE>
A:Cross-references: GB:X76989; NID:g475203; PIDN:CAA54294.1; PID:g475204
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
C:Keywords: glycosyltransferase; transmembrane protein

Alignment Scores:
```


436 GCCACAGAGGTGGGAATGAGATAGATCGATCTCTCGATTTGGGAAGATGACAAATGCC 493
 Db ::: |||||
 153 GluserSertyrGlyProGluLeuAspSerHisAspPheValLeuargMetAsnLysAla 172
 QY |||
 496 CCCACCAAGGTGTGAAGACATCGCGCGCATCCACCATGATTCGACGTGTCGCCCAT 555
 Db |||||
 173 ProThrAlaGlyPheGluAlaaspValGlyThrLysThrHisHisLeuValtyrPro 192
 QY |||||
 556 ACCACGCTTCCTCTTTGCTAAAAACCCCTGATTTATTTTCAAGGACGCAATACTACT 615
 Db |||
 193 Gluser-----PheArgGluLeuGlyAspAsn 201
 QY ::: |||||
 616 ATTGTGTTATTTCGGGACCTTTTCGCAATATCAGGAAGATGCCAATGCATCGTTTAC 675
 Db ::: |||||
 202 ValSerMetIleuValProPheLysThrIle-----AspLeuGluTrpValValSer 219
 QY ::: |||||
 676 AACATGTTGAAAGACACATGGT-----ATCATCCGAATGCCAAATATACGTG 726
 Db ::: |||||
 220 AlaIleThrThrGlyThrIleSerHisThrTyrlleProValProAlaLysIleargVal 239
 QY ::: |||||
 727 ACCACAGAGAACGCGCATGAGTTAC 750
 Db ::: |||||
 240 LysGlnAspLysIleLeuLeuItyr 247
 RESULT 10
 JC5251
 beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - human
 C:Species: Homo sapiens (man)
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 24-Sep-1999
 C:Accession: JC5251; G01021
 R:Kim, Y.J.; Kim, K.S.; Kim, S.H.; Kim, C.H.; Ko, J.H.; Choe, I.S.; Tsuji, S.;
 Biochem. Biophys. Res. Commun. 228, 324-327, 1996
 A:Title: Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,
 A:Reference number: JC5251; M01D:97079181; PMID:8920913
 A:Accession: JC5251
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-350 <KM>
 A:Cross-references: GB:U63090; NID:g1773282; PIDN:AA640389.1; PID:g1773283
 A:Experimental source: liver
 R:Giordanengo, V.
 Submitted to the EMBL Data Library, March 1996
 A:Reference number: H00561
 A:Accession: G01021
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-350 <GIO>
 A:Cross-references: EMBL:X96667; NID:g1235530
 C:Comment: This enzyme catalyzes the transfer of sialic acid from CMP-NeuAc to
 substrate preference for glycolipid than for O-linked oligosaccharides of glyco-
 C:Genetics:
 A:Gene: ST3(O)-II
 C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
 C:Keywords: glycosyltransferase

 Alignment Scores:
 Pred. No.: 0.00141 Length: 350
 Score: 131.50 Matches: 33
 Percent Similarity: 51.11% Conservative: 31
 Best Local Similarity: 36.67% Mismatches: 31
 Query Match: 5.52% Indels: 13
 DB: 2 Gaps: 4

 US-09-714-936-218 (1-1294) x JC5251 (1-350)
 QY 373 CAAGACGCTTTGCAACTGGACGTGTGACCTTTGGCCATAGTGTCAAACTCAGGTCAGATG 432
 Db ::: |||||
 144 ArgAspProHisGln-----CysArgArgCysAlaValValGlyAsnSerGlyAsnLeu 161
 QY |||
 433 GTTGGCCGAAGGTGGGAATGAGATAGATCGATCTCTCTCATTTGGGAATGAACAAT 492
 Db |||
 162 ArgGlySerGlyItyrGlnAspValAspGlyHisAsnPheIleMetArgMetAsnGln 181
 QY |||

```

QY 493 GCCCCACCAAGAGTGTATGAGAGATGTCGCCGCCATGACCATGATGAGTTCAGTGTGTGCC 552
Db 182 AlaProThrValGlyPheGluGlnAspValGlySer-----ArgThrHis 197
QY 553 CATACACAGGTCTCTCTTTTGCCTAAACCCCTGATTATTTTCAAGGAGCGAATACT 612
Db 198 HisPheMetTyrProGluSerAlaAlaLysAsn-----LeuProAlaAsnVal 212
QY 613 ACTATTTGTGTATTGGGACCTTTCCGC 642
Db 213 SerPheValLeuVal-----ProPheLys 220

RESULT 11
S36824
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1994 #sequence_revision 13-Mar-1997 #text_change 24-Sep-1999
C:Accession: S36824
R:Lee, Y.C.; Kurosawa, N.; Hamamoto, T.; Nakaoka, T.; Tsuji, S.
Eur. J. Biochem. 216, 377-385, 1993
A:Title: Molecular cloning and expression of Gal-beta-1,3GalNAc-alpha-2,3-sialyltransferase
A:Reference number: S36824; MUID:93387288; PMID:8375377
A:Status: preliminary
A:Accession: S36824
A:Molecule type: mRNA
A:Residues: 1-337 <LEP>
A:Cross-references: EMBL:X73523; NID:g402214; PIDN:CAA51919.1; PID:g402215
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
C:Keywords: glycosyltransferase

Alignment Scores:
Pred. No.: 0.00157 Length: 337
Score: 131.00 Matches: 50
Percent Similarity: 33.77% Conservative: 27
Best Local Similarity: 21.93% Mismatches: 57
Query Match: 5.50% Indels: 94
DB: 2 Gaps: 8

US-09-714-936-218 (1-1294) x S36824 (1-337)
QY 178 CTGAAGAGAAATCTGTGATGCTGTGAGCTTCATGACAGCGTTCCTTTTCCTGCTGTT 237
Db 1 MetArgArgLysThrLeuLysTyrLeuThrPhePheLeuLeuPheLeuLeu----- 18
QY 238 GTCCGCTGTGTAAGAAGTGAATTCCTTCCATTCCTACTAACTCCTTTGACACAACTGGT 297
Db 19 -----ThrSerPheValLeuAsnTyrSerAsnThrGlyValPro--- 31
QY 298 ACAAGTGGATACCA-----TTCTCTACATACAGGCGGCCCTT 339
Db 32 SerAlaTrpPheProLysGlnMetLeuLeuGluLeuSerGluAsnPheArgArgPheIle 51
QY 340 CGAAGCTCAC-----TATGGAATAC 357
Db 52 LysSerGlnProCysThrCysArgHisCysIleSerGlnAspLysValSerTyrTrpPhe 71
QY 358 -----ATAAAGTGAAGACACAGAGCCTTTG 384
Db 72 AspGlnArgPheAsn---LysThrMetGlnProLeuLeuThrValHisAsnAlaLeuMet 90
QY 384 ----- 384
Db 91 GluGluAspThrTyrArgTrpTrpLeuArgLeuGlnArgGluArgLysProAsnAsnLeu 110
QY 384 ----- 384
Db 111 SerAspThrValLysGluLeuPheArgLeuValProGlyAsnValAspProMetLeuAsn 130
QY 385 -----CAACTGGACTGTGACCTTTGTGTCATAGTGTCAAACTCAGTCAGTGTGGC 438
Db 131 LysArgLeuValGlyCysArgArgCysAlaValValGlyAsnSerGlyAsnLeuLysAsp 150

```

```

QY 439 CAGAAAGTGGGAATAGATAGATGATCCTCTGCAATTTGGAGATGAAATGCAATGCCCCC 498
Db 151 SerSerTyrGlyProGluLeuAspSerHisAspPheValLeuArgMetAsnLysAlaPro 170
QY 499 ACCAAAGCTTATGAAGATGTCGCCGCCATGACCATGATTCGAGTTGTGTCCTCATACC 558
Db 171 ThrValGlyPheGluAlaAspValGlySerArgThrHisHisLeuValTyrProGlu 190
QY 559 AGCGTCTCTCTTTTGTCTAAACCCCTGATTATTTTCAAGGAGCGAATACTACTATT 618
Db 191 Ser-----PheArgGluLeuGlyGluAsnVal 199
QY 619 TGTGTTATTTGGGACCTTTCCGC 642
Db 200 AsnMetValLeuValProPheLys 207

RESULT 12
JC6321
CMP-NeuAc-GM3 alpha2-8 sialyltransferase (EC 2.4.99.-) [imported] - rat
N:Alternate names: GD3-synthase; GM3-synthase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-May-1998 #sequence_revision 29-May-1998 #text_change 18-Aug-2000
C:Accession: JC6321
R:Zeng, G.; Gao, L.; Yu, R.K.
Gene 187, 131-134, 1997
A:Title: Cloning of the cDNA coding for rat brain CMP-NeuAc:CD3 alpha2-8 sialyltrans-
A:Reference number: JC6321; MUID:97225806; PMID:9073076
A:Accession: JC6321
A:Molecule type: mRNA
A:Residues: 1-380 <ZEN>
A:Cross-references: GB:U55938
A:Experimental source: brain
A:Comment: This enzyme is one of the sialyltransferases and adds one more alpha-2,8-
C:Superfamily: alpha-N-acetylneuraminate alpha-2,8-sialyltransferase
C:Keywords: glycosyltransferase
F:18-33/Domain: transmembrane #status predicted <TM>

Alignment Scores:
Pred. No.: 0.00451 Length: 380
Score: 126.00 Matches: 31
Percent Similarity: 47.52% Conservative: 17
Best Local Similarity: 30.69% Mismatches: 33
Query Match: 5.29% Indels: 20
DB: 2 Gaps: 3

US-09-714-936-218 (1-1294) x JC6321 (1-380)
QY 397 GACCTTTGGCCATGATGTCAAACTCAGTCAGATGGTGGCCAGAGTGGGAATGAG 456
Db 160 AsnIleCysAlaValValGlyAsnSerGlyIleLeuThrGlySerGlnCysGlyGlnGlu 179
QY 457 ATAGATCGATCCTCTCTGATTTGGAGATGACACATGCCCCCACCACCAAGTTATGAGAA 516
Db 180 IleAspLysSerAspPheValPheArgCysAsnPheAlaProThrGluAlaPheHisLys 199
QY 517 GATGTGGCGCCGATGACCATGATTCGAGTTGTGTGCCATACACCGCTTCTCTTTTGCTA 576
Db 200 AspValGlyLysLysThrAsnLeuThrPhe----- 210
QY 577 AAAACCCCTGATTATTTTCAAGGAGCGAATACTACTATTGTGTTATTTGGGACCT 636
Db 211 ---AsnProSerIleLeuGluLysTyrTyrAsnAsnLeuLeuThrIle----- 225
QY 637 TTCGCCATATGAGGAAGATGCGATGCGATCGCTTTACACATGTTGAAAGACAGTT 696
Db 226 -----GlnAspArgAsnAsnPhe----PheLeuSerLeuLysLysLeuAsp 239
QY 697 GGT 699
Db 240 Gly 240

RESULT 13
A45073

```

C; Keywords: transmembrane protein

```
Alignment Scores:
Pred. No.:          0.00686      Length:          380
Score:             124.00       Matches:         32
Percent Similarity: 46.53%     Conservative:    15
Best Local Similarity: 31.68%   Mismatches:     34
Query Match:       5.20%       Indels:         20
DB:                2           Gaps:            3

US-09-714-936-218 (1-1294) x A56950 (1-380)
QY 397 GACCTTGTGGCATTGTCACAACTCAGCTCAGATGTTGCCAGAGGTGGCAATGAG 456
Db 160 AsnValcysAlaValvalcylAsnSerGlylleLeuThrGlySerGlnCysGlyGlnGlu 179
QY 457 ATAGATCGATCCCTCTCGATTGTCAGATTGAGGAATGAACAATGCCCCACCACCAAGGTTATGAAGAA 516
Db 180 IleasPlysSerAspPheValserAArgCysAsnPheAlaProThrGluAlaPheHisLys 199
QY 517 GATGTGGCGCCGATCACCATCATGCATTCGATGTGTGCCATACACCGCGTTCCTCTTTTGCTA 576
Db 200 AspValglyArgLysThrAsnLeuThrPhrPe----- 210
QY 577 AAAAACCCCCGATTATTTTTTTCAGGAGGACGACTACTACTATTGTTGTTATTATTTGGGGAOCT 636
Db 211 --AsnProSerIleLeuGluLysTyrrAsnAsnLeuLeuThrlie----- 225
QY 637 TTCGCGCAATATGAGGAAGATGGCAATGGCAATGGCTTTACAACATGTTGAAAAGACAGCTT 696
Db 226 -----GlnAspArgAsnAsnPhe---PheLeuSerLeuLysLysLeuAsp 239
QY 697 GGT 699
Db 240 Gly 240

RESULT 15
G86216
protein T23G18.14 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G86216
R;neologids: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; D.
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;
C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Mailli, R.;
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MWID:21016719; PMID:11130712
A;Accession: G86216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <F>O>
A;Cross-references: GS:AEO05172; NID:96579198; PIDN:NAF18241.1; GSPDB:GN00141
C;Genetics:
A;Gene: T23G18.14
A;Map position: 1

Alignment Scores:
Pred. No.:          0.00848      Length:          398
Score:             123.00       Matches:         49
Percent Similarity: 40.09%     Conservative:    36
Best Local Similarity: 23.11%   Mismatches:     82
Query Match:       5.16%       Indels:         45
DB:                2           Gaps:            7

US-09-714-936-218 (1-1294) x G86216 (1-398)
QY 9 TC CGGGTCGACGATTCGCTCGCGGTGCCCTCTATTATTTGGATTCGGGAATGTGGGTGGA 68
```

```

    |||    :::    |||||:::    |||||:::|:::|
30 SerSerPheAlaSerAlaIleAlaGluLeuGlySerSerdGlyLeuMetThrGlu 49
QY 69 GAGGTCTCTCCGTGTACAGCTCCAGCTGCCGCCAGGACTGCCCTGACCCAGGCGC 128
Db 50 AspIleVal-PheAsnGluThrLeu-----LeuGluPheAlaIleAspProGlyCl 67
QY 129 GCCC-----CCTCTCGGTGGCAGGAGGCGCGCGGAGCGC 164
Db 67 uProAsnPheLysGlnGluValAspLeuIleSerAspTyrAspHisThrArgArgSer-- 86
QY 165 CATGGCCTCCACTGAGCAAGTCTGTGTTCTGTGAGCTTCATACGACGCTTCT 224
Db 87 -----HisArgArgHisPheSerSerMetSerIleArgProSerGluCl 101
QY 225 TTTCCTGCTGGTGTGCGT---CTGTAAATGAAGTGAATTCCTACTTACTAACTG 281
Db 101 nGlnArgValSerArgAspIleAlaSerSerLysPheProValThrLeuArgSe 121
QY 282 CTTTGCACAACTGCTACAAAGTGTACCATTCCTCCTACACATACAGG----- 330
Db 121 r---SerGlnAlaTyrArgTyrTrpSerGluPheLysArgAsnLeuArgLeuTrpAlaAr 140
QY 331 -----CGGCCCT 338
Db 140 gArgArgAlaTyrGluProAsnIleMetLeuAspLeuIleArgLeuValLysAsnProIl 160
QY 339 TCGAACTCACTATGGATACATAAATGTGAGACACACAGAGCCTTTCGAAGTGTGA 398
Db 160 eAspValHisAsnGlyValValSerIleSerSerGluArgTyrLeuSer----- 176
QY 399 CCTTTGTGCCATAGTGTCAACTCAGTCAGTCAGTGTGGCCAGAGGTGGGAAATCAGAT 458
Db 177 ----CysAlaValValGlyAsnSerGlyThrLeuLeuAsnSerGlnTyrGlyAspLeuIl 195
QY 459 AGATPCATCTCTCCTCATTTGGAGAAATGAACAATCCCCCACCACCAAGGTATGAAGAAGA 518
Db 195 eAspLysHisGluIleValIleArgLeuAsnAlaLysThrGluArgPheGluLysLy 215
QY 519 TGTCGGCCCATGACCATGATTCGAGTTGTGCC 552
Db 215 sValGlySerLysThrAsnIleSerPheIleAsn 226
```

Search completed: September 12, 2003, 16:37:38
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 12, 2003, 16:37:44 : Search time 52 Seconds
(without alignments)
7261.949 Million cell updates/sec

Title: US-09-714-936-218
Perfect score: 2383
Sequence: 1 ccggaattccgggtgcagc.....ttttctctctttttttg 1294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 1083872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPro.spool/US09714936/runat_10092003_151503_5912/app_query.fasta.1.1479
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MTN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09714936.cgn_1.1.53 @runat_10092003_151503_5912
-NCPU=6 -ICPU=3 -NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Query Match	Score	Length	DB ID	Description

1	1095	46.0	210	12	US-10-199-672-536	Sequence 536, App
2	1095	46.0	210	12	US-10-187-749-536	Sequence 536, App
3	1095	46.0	210	12	US-10-194-457-536	Sequence 536, App
4	1095	46.0	210	12	US-10-184-642-536	Sequence 536, App
5	1095	46.0	210	12	US-10-196-747-536	Sequence 536, App
6	1095	46.0	210	12	US-10-173-689-536	Sequence 536, App
7	1095	46.0	210	12	US-10-173-890-536	Sequence 536, App
8	1095	46.0	210	12	US-10-173-891-536	Sequence 536, App
9	1095	46.0	210	12	US-10-173-692-536	Sequence 536, App
10	1095	46.0	210	12	US-10-173-694-536	Sequence 536, App
11	1095	46.0	210	12	US-10-173-698-536	Sequence 536, App
12	1095	46.0	210	12	US-10-173-699-536	Sequence 536, App
13	1095	46.0	210	12	US-10-173-707-536	Sequence 536, App
14	1095	46.0	210	12	US-10-174-569-536	Sequence 536, App
15	1095	46.0	210	12	US-10-174-583-536	Sequence 536, App
16	1095	46.0	210	12	US-10-174-587-536	Sequence 536, App
17	1095	46.0	210	12	US-10-174-589-536	Sequence 536, App
18	1095	46.0	210	12	US-10-174-591-536	Sequence 536, App
19	1095	46.0	210	12	US-10-175-736-536	Sequence 536, App
20	1095	46.0	210	12	US-10-175-742-536	Sequence 536, App
21	1095	46.0	210	12	US-10-175-744-536	Sequence 536, App
22	1095	46.0	210	12	US-10-175-745-536	Sequence 536, App
23	1095	46.0	210	12	US-10-175-748-536	Sequence 536, App
24	1095	46.0	210	12	US-10-175-751-536	Sequence 536, App
25	1095	46.0	210	12	US-10-175-754-536	Sequence 536, App
26	1095	46.0	210	12	US-10-176-480-536	Sequence 536, App
27	1095	46.0	210	12	US-10-176-489-536	Sequence 536, App
28	1095	46.0	210	12	US-10-176-754-536	Sequence 536, App
29	1095	46.0	210	12	US-10-176-755-536	Sequence 536, App
30	1095	46.0	210	12	US-10-176-759-536	Sequence 536, App
31	1095	46.0	210	12	US-10-176-920-536	Sequence 536, App
32	1095	46.0	210	12	US-10-176-922-536	Sequence 536, App
33	1095	46.0	210	12	US-10-176-924-536	Sequence 536, App
34	1095	46.0	210	12	US-10-176-984-536	Sequence 536, App
35	1095	46.0	210	12	US-10-179-508-536	Sequence 536, App
36	1095	46.0	210	12	US-10-179-512-536	Sequence 536, App
37	1095	46.0	210	12	US-10-179-515-536	Sequence 536, App
38	1095	46.0	210	14	US-10-052-586-536	Sequence 536, App
39	1095	46.0	210	15	US-10-174-590-536	Sequence 536, App
40	1095	46.0	210	15	US-10-176-758-536	Sequence 536, App
41	1095	46.0	210	15	US-10-175-737-536	Sequence 536, App
42	1095	46.0	210	15	US-10-173-706-536	Sequence 536, App
43	1095	46.0	210	15	US-10-175-738-536	Sequence 536, App
44	1095	46.0	210	15	US-10-175-752-536	Sequence 536, App
45	1095	46.0	210	15	US-10-176-482-536	Sequence 536, App

ALIGNMENTS

RESULT 1
US-10-199-672-536
; Sequence 536, Application US/10199672
; Publication No. US20030148442A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/199, 672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263

```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-536

Alignment Scores:
Pred. No.:      1,03e-101      Length:      210
Score:          1095.00      Matches:      206
Percent Similarity: 99.52%      Conservative: 0
Best Local Similarity: 99.52%      Mismatches: 1
Query Match:      45.95%      Indels:      0
DB:              12      Gaps:      0

US-09-714-936-218 (1-1294) x US-10-199-672-536 (1-210)
QY 166 AFGCGCTGCATCTCTGAGAGAAAGTCTGTGATTCGTGTGAGCTTCATACAGCGTTCCTT 225
Db 1 MetAlaCysIleuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGGTGGTGGTCTTCTTAATCAAGTGAATTCCTCACTACTAACTGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40
QY 286 GSACAACCTGGTCAAAAGTGATACCATCTCTACATACAGCGGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysrPileProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGGATACATAATGTGAGACACAAAGCGCTTGCACACTGGACTGTGACCTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysPheCys 80
QY 406 GCATAGTGTCAAACTCAGGTGAGTGTGGCCAGAGAGTGGGAAATGAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCTGCTGATTCGAGATGAAATGCAATGCCCCACCAAGGTTATGAGAGATGTCGCG 525
Db 101 SerSerCysIleThrPargMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CGCATGACCATGATTCGAGTGTGTCGCCATACGAGCGTTCCTCTTTCGCTAAACACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCAATATCTACTATTGTGTTATTTGGGACCTTCCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleThrPheProPheArgAsn 160
QY 646 ATGAGGAAGATGCGCAATGCGATTCGCTTTACACATGTTGAAAGACAGCTGGATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGATGCGCCAAATATACGTGACCCACAGAGAGCGCATGATGCTGATGAGTGTCTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGGAACACTGGGAAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207

RESULT 2
US-10-187-749-536
; Sequence 536, Application US/10187749
; Publication No. US20030153036A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/187,749
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-749-536

Alignment Scores:
Pred. No.:      1,03e-101      Length:      210
Score:          1095.00      Matches:      206
Percent Similarity: 99.52%      Conservative: 0
Best Local Similarity: 99.52%      Mismatches: 1
Query Match:      45.95%      Indels:      0
DB:              12      Gaps:      0

US-09-714-936-218 (1-1294) x US-10-187-749-536 (1-210)
QY 166 AFGCGCTGCATCTCTGAGAGAAAGTCTGTGATTCGTGTGAGCTTCATACAGCGTTCCTT 225
Db 1 MetAlaCysIleuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGGTGGTGGTCTTCTTAATCAAGTGAATTCCTCACTACTAACTGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuAsnCysPhe 40
QY 286 GSACAACCTGGTCAAAAGTGATACCATCTCTACATACAGCGGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysrPileProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGGATACATAATGTGAGACACAAAGCGCTTGCACACTGGACTGTGACCTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysPheCys 80
QY 406 GCATAGTGTCAAACTCAGGTGAGTGTGGCCAGAGAGTGGGAAATGAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCTGCTGATTCGAGATGAAATGCAATGCCCCACCAAGGTTATGAGAGATGTCGCG 525
Db 101 SerSerCysIleThrPargMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CGCATGACCATGATTCGAGTGTGTCGCCATACGAGCGTTCCTCTTTCGCTAAACACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCAATATCTACTATTGTGTTATTTGGGACCTTCCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleThrPheProPheArgAsn 160
QY 646 ATGAGGAAGATGCGCAATGCGATTCGCTTTACACATGTTGAAAGACAGCTGGATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGATGCGCCAAATATACGTGACCCACAGAGAGCGCATGATGCTGATGAGTGTCTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
```

QY 286 GGACAACTGGTACAAAGTGATACCACTCTCTACACATACAGCGGCCCTTCGAACT 345
 DB |||||||
 QY 41 GYGlnProGlyThrLysTrpIleProPheSerTyrThrArgProLeuArgThr 60
 DB |||||||
 QY 346 CACTATGGATACATAAATGTGAAGACACAGAGCCCTTTGCAACTGGAGCTGCACTTTGT 405
 DB |||||||
 QY 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
 DB |||||||
 QY 406 GCCATAGTCTCAACTCAGTTCAGTGGTGGCCAGAGGTGGGAAATGATGATCGA 465
 DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
 QY 466 TCCTCTGCTGATTTGGAGAAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTCGGC 525
 DB 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTyrGlnGluAspValGly 120
 QY 526 CGCATGACCATGATTCGAGTGTGTCCCATACACAGCGTTCCTCTTTGCTTAAACACCT 585
 DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
 QY 586 GATTATTTTTCAGGAGCGCAATACACTACTACTATTTTGTGTATTTGGGACCTTCCCGCAAT 645
 DB 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
 QY 646 ATGAGAAAGAGATGGCAATGCGCTTTACAAACATGTTGAAAAAGACAGTTGATCTAT 705
 DB 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
 QY 706 CCGAATGCCAAATATAGTGGACACAGAGAGCCCATGAGTTACTGTGATGAGTTT 765
 DB 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAAGGAAGTGGGAAGGAC 786
 DB 201 LysLysGluThrGlyLysAsp 207

RESULT 3
 US-10-194-457-536
 ; Sequence 536 Application US/10194457
 ; Publication No. US20030153037A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430RIC296
 ; CURRENT APPLICATION NUMBER: US/10/194,457
 ; CURRENT FILING DATE: 2002-07-11
 ; PRIOR APPLICATION NUMBER: 10/052586
 ; PRIOR FILING DATE: 2002-01-15
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059266
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063120
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063121
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063486
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/063540
 ; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063541
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063544
 ; PRIOR FILING DATE: 1997-10-28
 ; Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 536
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-194-457-536

Alignment Scores:
 Pred. No.: 1.03e-101 Length: 210
 Score: 1095.00 Matches: 206
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 45.95% Indels: 0
 DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-194-457-536 (1-210)

QY 166 ATGSCCTGCATCCTCGAAGAAAGTCTGTGATTTGTGTGAGCTTCATAGACAGCGTTCCCTT 225
 DB 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaIlePheLeu 20
 QY 226 TTCTCTGCTGTTGCGCTGTTGTAATGAAGTGAATTTCCCATTTGCTACTAACTGCTTT 285
 DB 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
 QY 286 GGACAACTGGTACAAAGTGATACCACTCTCTACACATACAGCGGCCCTTCGAACT 345
 DB 41 GYGlnProGlyThrLysTrpIleProPheSerTyrThrArgProLeuArgThr 60
 QY 346 CACTATGGATACATAAATGTGAAGACACAGAGCCCTTTGCTTAAACACCTTTGT 405
 DB 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
 QY 406 GCCATAGTCTCAACTCAGTTCAGTGGTGGCCAGAGGTGGGAAATGATGATCGA 465
 DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
 QY 466 TCCTCTGCTGATTTGGAGAAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTCGGC 525
 DB 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTyrGlnGluAspValGly 120
 QY 526 CGCATGACCATGATTCGAGTGTGTCCCATACACAGCGTTCCTCTTTGCTTAAACACCT 585
 DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
 QY 586 GATTATTTTTCAGGAGCGCAATACACTACTACTATTTTGTGTATTTGGGACCTTCCCGCAAT 645
 DB 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
 QY 646 ATGAGAAAGAGATGGCAATGCGCTTTACAAACATGTTGAAAAAGACAGTTGATCTAT 705
 DB 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
 QY 706 CCGAATGCCAAATATAGTGGACACAGAGAGCCCATGAGTTACTGTGATGAGTTT 765
 DB 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAAGGAAGTGGGAAGGAC 786
 DB 201 LysLysGluThrGlyLysAsp 207

RESULT 4
 US-10-184-642-536
 ; Sequence 536 Application US/10184642
 ; Publication No. US20030157635A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian

QY 346 CACTATGATACATAAAATCTGAACACAGAGAGCTTTGCAACTGGAGTGTGACCTTTGT 405
Db 61 HistyrglytyrileasnvallythrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
QY 406 GCCTATGTCCTGCAACTGAGTCAGATGGTGGCCAGAGGTGGGAATGAGATAGATCGA 465
Db 81 AlailevalserasnsersglylnmetvalglylnLysvalglyAsnGluileAspArg 100
QY 466 TCCCTCTGCTGCTGAGAGTGAACAATGCCCCCACCACAAAGGTTATGAAGAAGATGTCGCG 525
Db 101 SerSerCysiletrpArgMetasnsalaproThrLysGlytyrGluLusaspValGly 120
QY 526 CGCATGACCATGATTCGAGTGTGTCCCATACAGAGGTGTCCTTTGCTAAAAACCT 585
Db 121 ArgMetThrMetileargValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATGACTACTATTGTTGTTATTGTTGGGACCTTCCGCAAT 645
Db 141 AspyrPhePhelyGluAlaasnThrThrIleTyrrValiletrpGlyProPheargasn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACACATCTTGAAAGACAGAGTGTGATCTAT 705
Db 161 MetArgLysaspLysnGlyileValTyrrasnMetLeuLysLysThrValGlyIleTy 180
QY 706 CCGAATCCCAATATAGTGCACACAGAGAGCGCATGAGTACTGTGTGATGGAGTTT 765
Db 181 ProAsnAlaGlnIleTyrrValThrThrGluLysArgMetSerTyrcysaspGlyValPhe 200
QY 766 AAGAAGAAACTGGGAAGAC 786
Db 201 LysLysGluThrGlyLysasp 207

RESULT 8

US-10-173-691-536

; Sequence 536, Application US/10173691

; Publication No. US20030166106A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C13

; CURRENT APPLICATION NUMBER: US/10/173,691

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 536

; LENGTH: 210

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-691-536

Alignment Scores:

Pred. No.:	1,03e-101	Length:	210
Score:	1095.00	Matches:	206
Percent Similarity:	99.52%	Conservative:	0
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	45.95%	Indels:	0
DB:	12	Gaps:	0

US-09-714-936-218 (1-1284) x US-10-173-691-536 (1-210)

QY 166 ATGGCCTGATCTCTGAGAGAGAGTCTGTGATTCTGTGAGCTTCTATGACAGCGTTCCTT 225

Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCCTGCTGGTGGCCCTTGTGAATGAAGTGAATTTCCCATTTGCTACTAACTGCTTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACCTGGTACAAAAGTGGATPACCATTCCTACACATACAGCGCGCCCTTGGAACT 345
Db 41 GlyGlnProGlyThrLysTyrPheProPheSerTyrrThrTyrrArgProLeuArgThr 60
QY 346 CACTATGATACATAAAATCTGAAGACACAGAGCGCTTGGCACTGAGCTGTGACCTTTGT 405
Db 61 HistyrglytyrileasnvallythrGlnGluProLeuGlnLeuAspCysAspLeuCys 80
QY 406 GCCATAGTGTCAAACTCAGATGCTGAGTGGTGGCCAGAGGTGGGAATGAGATAGATCGA 465
Db 81 AlailevalserasnsersglylnmetvalglylnLysvalglyAsnGluileAspArg 100
QY 466 TCCCTCTGATTTGGGAATGAACAATGCCCCCACCACAAAGTGTATGAAGAAGATGTCGCG 525
Db 101 SerSerCysiletrpArgMetasnsalaproThrLysGlytyrGluLusaspValGly 120
QY 526 CCGATGACCATGATTCGAGTGTGTCCCATACAGCGTTCCTTTGCTAAAAACCT 585
Db 121 ArgMetThrMetileargValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATGACTACTATTGTTGTTATTGTTGGGACCTTCCGCAAT 645
Db 141 AspyrPhePhelyGluAlaasnThrThrIleTyrrValiletrpGlyProPheargasn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACAACTGTGTGAAAGACAGAGTGTGATCTAT 705
Db 161 MetArgLysaspLysnGlyileValTyrrasnMetLeuLysLysThrValGlyIleTy 180
QY 706 CCGAATCCCAATATAGTGCACACAGAGAGCGCATGAGTACTGTGTGATGGAGTTT 765
Db 181 ProAsnAlaGlnIleTyrrValThrThrGluLysArgMetSerTyrcysaspGlyValPhe 200
QY 766 AAGAAGAAACTGGGAAGAC 786
Db 201 LysLysGluThrGlyLysasp 207

RESULT 9

US-10-173-692-536

; Sequence 536, Application US/10173692

; Publication No. US20030166188A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C20

; CURRENT APPLICATION NUMBER: US/10/173,692

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 536

; LENGTH: 210

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-692-536

Alignment Scores:

Pred. No.: 1.03e-101 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservatives: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-173-692-536 (1-210)

```
QY 166 ATGGCTGCATCTCTGAGAAAGTCTGTGATTCCTGAGCTTCATAGCAGCTTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGCTGCTCTGCTCTGTAATGAAGTGAATTCCTCATCTACTAACTGCTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACTGTCATAAGTGGATACCATTCCTACATACAGAGCGGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGGAAGACACAGAGCCTTTCGCACTGGACTGGACCTTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysLeuCys 80
QY 406 GCATAGTGTCAAACTCAGTCTGATGTTGCCAGAGGTGGGAAATGAGATAGATCGA 455
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTTGGAGATGAACATGCCCCACCAAGGTTATGAAGAAGATGCGC 525
Db 101 SerSerCysIleTrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CCATGACCAATTCAGTGTGTGCCATACAGAGCTTCTCTTTTCTATAAAACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTATTGTTGGGACCTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGAAAGATGCAATGCGATCGTTTACAACTGTTGAAAGACAGATTTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGAATGCCAAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGCTTTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGAACTGGGAAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207
```

RESULT 10

US-10-173-694-536

; Sequence 536, Application US/10173694

; Publication No. US20030166107A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P34301C19

; CURRENT APPLICATION NUMBER: US/10/173,694

; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-694-536

Alignment Scores:

Pred. No.: 1.03e-101 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservatives: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-173-694-536 (1-210)

```
QY 166 ARGCCCTGCATCTCTGAGAAAGTCTGTGATTCCTGAGCTTCATAGCAGCTTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaPheLeu 20
QY 226 TTCCTGCTGCTGCTCTGCTCTGTAATGAAGTGAATTCCTCATCTACTAACTGCTT 285
Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACTGTCATAAGTGGATACCATTCCTACATACAGAGCGGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGGAAGACACAGAGCCTTTCGCACTGGACTGGACCTTTGT 405
Db 61 HisTyrGlyTyrIleAsnValLysThrGlnGluProLeuGlnLeuAspCysLeuCys 80
QY 406 GCATAGTGTCAAACTCAGTCTGATGTTGCCAGAGGTGGGAAATGAGATAGATCGA 465
Db 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCCTGCATTTGGAGATGAACATGCCCCACCAAGGTTATGAAGAAGATGCGC 525
Db 101 SerSerCysIleTrpArgMetAsnAsnAlaProThrLysGlyTyrGluGluAspValGly 120
QY 526 CCATGACCAATTCAGTGTGTGCCATACAGAGCTTCTCTTTTCTATAAAACCT 585
Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCGAATACTACTATTGTGTATTGTTGGGACCTTCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGAAAGATGCAATGCGATCGTTTACAACTGTTGAAAGACAGATTTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGAATGCCAAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGCTTTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGAACTGGGAAGGAC 786
Db 201 LysLysGluThrGlyLysAsp 207
```

RESULT 11

US-10-173-698-536

; Sequence 536, Application US/10173698

; Publication No. US20030166108A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

```

; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1G12
; CURRENT APPLICATION NUMBER: US/10/173,698
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-698-536

Alignment Scores:
Pred. No.: 1,03e-101 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-173-698-536 (1-210)
QY 166 ATGGCGTCATCTCTGAAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGCGTTCCCTT 225
DB 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
QY 226 TTCCTGCTGGTGTGGCGTCTTGAATGAATGAAGTAATTCCTGCTACTAACTGCTTTT 285
DB 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACCTGTGCAAAAGTGATACCATTCCTCTACATACAGCGCGCCCTTCGAACT 345
DB 41 GlyGlnProGlyThrLysTrpIleProPheSerThrLysArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTGAAGACACACAGAGCGCTTCGCAAGGTATGAAGAAGATGTCGC 525
DB 61 HistyrGlyTrpIleAsnValLysThrGlnLysValGlyAsnGluLysParg 100
QY 406 GCATATGTCATCAAAATGTGAAGACACACAGAGCGCTTCGCAAGGTATGAAGAAGATGTCGC 585
DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluLysParg 120
QY 466 TCCCTCCGCAATTTGGAGATGAACAATGCCCCACCAAGGTATGAAGAAGATGTCGC 705
DB 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTrpGluLysParg 140
QY 526 CGCANGACCATGATCGAGTTGTGCCCATACAGCGTTCCTCTTTTGTAAAAACCCCT 645
DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 160
QY 586 GATTATTTTTCAGAGAGCGCAATACCTATTATTGTTGTTATTGGGACCTTTCCGCAT 765
DB 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 180
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACATGTTGAAAAGACAGTGGTATCTAT 705
DB 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 190
QY 706 CGCAATGCCCAATATACGPGACACACAGAGAGCGCATGAGTTACGTGATGGAGTTT 765
DB 181 ProAsnAlaGlnIleTyrValThrThrGlnLysArgMetSerTyrCysAspValPhe 200
QY 766 AAGAAGGAACCTGGGAAGGAC 786
DB 201 LysLysGluThrGlyLysAsp 207

```

```

RESULT 12
US-10-173-699-536
; Sequence 536, Application US/10173699
; Publication No. US20030166109A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C8
; CURRENT APPLICATION NUMBER: US/10/173,699
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-699-536

Alignment Scores:
Pred. No.: 1,03e-101 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-173-699-536 (1-210)
QY 166 ATGGCGTCATCTCTGAAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGCGTTCCCTT 225
DB 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20
QY 226 TTCCTGCTGGTGTGGCGTCTTGAATGAAGTAATTCCTGCTACTAACTGCTTTT 285
DB 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACCTGTGCAAAAGTGATACCATTCCTCTACATACAGCGCGCCCTTCGAACT 345
DB 41 GlyGlnProGlyThrLysTrpIleProPheSerThrLysArgArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTGAAGACACACAGAGCGCTTCGCAAGGTATGAAGAAGATGTCGC 405
DB 61 HistyrGlyTrpIleAsnValLysThrGlnLysValGlyAsnGluLysParg 80
QY 406 GCATATGTCATCAAAATGTGAAGACACACAGAGCGCTTCGCAAGGTATGAAGAAGATGTCGC 465
DB 81 AlaIleValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluLysParg 100
QY 466 TCCCTCCGCAATTTGGAGATGAACAATGCCCCACCAAGGTATGAAGAAGATGTCGC 525
DB 101 SerSerCysIleTrpArgMetAsnAlaProThrLysGlyTrpGluLysParg 120
QY 526 CGCANGACCATGATCGAGTTGTGCCCATACAGCGTTCCTCTTTTGTAAAAACCCCT 585
DB 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGAGAGCGCAATACCTATTATTGTTGTTATTGGGACCTTTCCGCAT 645
DB 141 AspTyrPhePheLysGluAlaAsnThrThrIleTyrValIleTrpGlyProPheArgAsn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACATGTTGAAAAGACAGATGGTATCTAT 705

```

Db 161 MetArgLysAspGlyAsnGlyLeValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
 QY 706 CCGATGCCCAATATACGTGACACAGAGCGCATGAGTTACTGTGATGAGTTTTT 765
 Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAAGGAAGTGGGAAGGAC 786
 Db 201 LysLysGluThrGlyLysAsp 207

RESULT 13
 US-10-173-707-536
 ; Sequence 536, Application US/10173707
 ; Publication No. US20030166110A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430RIC17
 ; CURRENT APPLICATION NUMBER: US/10/173.707
 ; CURRENT FILING DATE: 2002-06-17
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 536
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-173-707-536

Alignment Scores:
 Pred. No.: 1,03e-101 Length: 210
 Score: 1095.00 Matches: 206
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 45.95% Indels: 0
 DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-173-707-536 (1-210)

QY 526 CGCATGACCATGATTCGAGTTGTCCTCCCATACAGCGTTCCTCTTTGCTAAAAACCT 585
 Db 121 ArgMetThrMetIleArgValValSerHisThrSerValProLeuLeuLysAsnPro 140
 QY 586 GATTATTTTTCAGGAGGAAATACTACTATTGTTGTTATTTGGGACCTTTCGCAAT 645
 Db 141 AspTyrPhePheLysGluAlaAsnThrThIleTyrValIleTyrGlyProPheArgAsn 160
 QY 646 ATGAGGAAAGATGGCAATGGCATCGTTTACACATGTTGAAAAAGACAGTGGTATCTAT 705
 Db 161 MetArgLysAspGlyAsnGlyLeValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
 QY 706 CCGATGCCCAATATACGTGACACAGAGCGCATGAGTTACTGTGATGAGTTTTT 765
 Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
 QY 766 AAGAAGGAAGTGGGAAGGAC 786
 Db 201 LysLysGluThrGlyLysAsp 207

RESULT 14
 US-10-174-569-536
 ; Sequence 536, Application US/10174569
 ; Publication No. US20030166111A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430RIC39
 ; CURRENT APPLICATION NUMBER: US/10/174.569
 ; CURRENT FILING DATE: 2002-06-18
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 536
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-174-569-536

Alignment Scores:
 Pred. No.: 1,03e-101 Length: 210
 Score: 1095.00 Matches: 206
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 45.95% Indels: 0
 DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-174-569-536 (1-210)

QY 166 ATGCGCTGCATCCTGAAGAGAAAGTCTGTGATTCGTGTGAGCTTCATAGCAGCGTTCCCTT 225
 Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaLaPheLeu 20
 QY 226 TTCTCTCTGGTTCGGCTTGTAAATGAAGTCAATTTCCCATTCCTACTAAACCTGCTTT 285
 Db 21 PheLeuLeuValValArgLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
 QY 286 GGCAACACCTGGTACAAAGTGGATACCATTCCTCCTACATACAGCGCGCCCTTCGAAT 345
 Db 41 GlyGlnProGlyThrLysTrpIleProPheSerTyrThrTyrArgArgProLeuArgThr 60
 QY 346 CACTATGGATACATAAATGTGGAAGACACAGAGCCCTTTGCCAAGTGGACTGTGACCTTTGT 405

Db 61 HistyrGlytyrIleasnVallysthrGlnGluProLeuGlnLeuaspCysAspLeuCys 80
QY 406 GCCATAGTCTCAAACTCAGGTCAGATGGTGGCCAGAAAGTGGGAAATGAGATAGATCGA 465
Db 81 AlaileValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCTCGATTGGAGATGAACAATGCCCCACCAAGGTTATGAGAAGATCTCGCG 525
Db 101 SerSerCysIleTrrpArgMetAsnAlaProthrLysGlytyrGluGluaspValGly 120
QY 526 CGCATGACATGATTCGAGTGTGCCATACACGCTTCTCTTTTGTCTTAAAGAACCT 585
Db 121 ArgMetThrMetIleArgValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCAATFACTATTATTGTGTATTATTTGGGACCTTCCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaasnThrThrIleTyrValIleTrrpGlyPropheargAsn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACATGTTGAAAGACAGCTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGAATGCCCAATATAGTCGACACAGAGAACGCGCATGAGTTACTGTGATGAGTATTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGAACTGGGAAGAC 786
Db 201 LysLysGluThrGlyLysAsp 207

RESULT 15

US-10-174-583-536
; Sequence 536, Application US/10174583
; Publication No. US20030166112A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC22
; CURRENT APPLICATION NUMBER: US/10/174,583
; PRIOR FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 536
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-583-536

Alignment Scores:
Pred. No.: 1,036-101 Length: 210
Score: 1095.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 45.95% Indels: 0
DB: 12 Gaps: 0

US-09-714-936-218 (1-1294) x US-10-174-583-536 (1-210)

QY 166 ATGGCCCTGCATCTGAAGAAAGTGTGATTGCTGTGAGCTTCATAGACGCGTTCCTT 225
Db 1 MetAlaCysIleLeuLysArgLysSerValIleAlaValSerPheIleAlaAlaPheLeu 20

QY 226 TTCCTGCTGGTGTGGCTCTTGTAAATGAAGTGAATTCCTACTTCTACTAACTGCTTT 285
Db 21 PheLeuLeuValValargLeuValAsnGluValAsnPheProLeuLeuLeuAsnCysPhe 40
QY 286 GGACAACTCGTACAAAGTGGATACCAITTCCTACACATACAGAGCGGCCCTTCGAACT 345
Db 41 GlyGlnProGlyThrLysTrrpIleProPheSerTyrThrTrrArgProLeuArgThr 60
QY 346 CACTATGATACATAAATGTGAAGACACAGAGCCTTTCGCAACTGCACTGTGACCTTTGT 405
Db 61 HistyrGlytyrIleasnVallysthrGlnGluProLeuGlnLeuaspCysAspLeuCys 80
QY 406 GCCATAGTCTCAAACTCAGATGGTGGCCAGAAAGTGGGAAATGAGATAGATCGA 465
Db 81 AlaileValSerAsnSerGlyGlnMetValGlyGlnLysValGlyAsnGluIleAspArg 100
QY 466 TCCTCTCTCATTTGGGAATGAACATGCCCCACCAAGGTTATGAGAAGATCTCGCG 525
Db 101 SerSerCysIleTrrpArgMetAsnAlaProthrLysGlytyrGluGluaspValGly 120
QY 526 CGCATGACATGATTCGAGTGTGCCATACACGCTTCTCTTTTGTCTTAAAGAACCT 585
Db 121 ArgMetThrMetIleArgValSerHisThrSerValProLeuLeuLeuLysAsnPro 140
QY 586 GATTATTTTTCAGGAAGCAATFACTATTATTGTGTATTATTTGGGACCTTCCGCAAT 645
Db 141 AspTyrPhePheLysGluAlaasnThrThrIleTyrValIleTrrpGlyPropheargAsn 160
QY 646 ATGAGGAAGATGGCAATGGCATCGTTTACACATGTTGAAAGACAGCTGGTATCTAT 705
Db 161 MetArgLysAspGlyAsnGlyIleValTyrAsnMetLeuLysLysThrValGlyIleTyr 180
QY 706 CGGAATGCCCAATATAGTCGACACAGAGAACGCGCATGAGTTACTGTGATGAGTATTT 765
Db 181 ProAsnAlaGlnIleTyrValThrThrGluLysArgMetSerTyrCysAspGlyValPhe 200
QY 766 AAGAAGAACTGGGAAGAC 786
Db 201 LysLysGluThrGlyLysAsp 207

Search completed: September 12, 2003, 16:50:26
Job time : 86 secs

1	98	7.6	1556	4	US-09-599-3608-25	Sequence 25, Appl
2	88.4	6.8	2056	3	US-09-334-601-12	Sequence 12, Appl
3	82.6	6.4	2697	4	US-09-620-312D-401	Sequence 401, App
4	50.6	3.9	1020	2	US-08-446-875-15	Sequence 15, Appl
5	49.8	3.8	1218	2	US-08-446-875-1	Sequence 1, Appl
6	49.8	3.8	1218	2	US-08-102-885-1	Sequence 1, Appl
7	36.8	2.8	1919	1	US-07-991-887A-1	Sequence 1, Appl
8	36.8	2.8	1919	1	US-08-309-985-1	Sequence 1, Appl
9	36.5	2.8	1048	1	US-08-626-994A-4	Sequence 4, Appl
10	36.6	2.8	1048	3	US-08-957-742-4	Sequence 4, Appl
11	36.6	2.8	1660	1	US-08-626-994B-2	Sequence 2, Appl
12	36.6	2.8	1660	3	US-08-957-742-2	Sequence 2, Appl
13	35.6	2.8	2528	4	US-09-596-243-346	Sequence 346, App
14	35.4	2.7	319608	4	US-09-539-333D-1	Sequence 1, Appl
15	35.4	2.7	319608	4	US-09-679-409-1	Sequence 1, Appl
16	34	2.6	1158	2	US-08-446-875-11	Sequence 11, Appl
17	34	2.6	1158	2	US-08-102-885B-11	Sequence 11, Appl
18	34	2.6	1766	1	US-07-991-887B-6	Sequence 6, Appl
19	34	2.6	1766	1	US-08-309-985-6	Sequence 6, Appl
20	33.8	2.6	636	4	US-09-252-991A-1494	Sequence 1494, A
21	33.8	2.6	816	4	US-09-252-991A-14751	Sequence 14751, A
22	33.8	2.6	1605	4	US-09-253-991A-14869	Sequence 14869, A
23	33.6	2.6	1665	4	US-09-247-155-72	Sequence 72, Appl
24	33.6	2.6	1666	2	US-08-666-367B-3	Sequence 3, Appl
25	33.4	2.6	1666	3	US-08-143-438-3	Sequence 3, Appl
26	33.4	2.6	7218	1	US-08-233-463-14	Sequence 14, Appl
27	33.4	2.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl


```
Db 744 GGCTGGTGTCTCCCAACATGGAAGCATATGCGTCTCTCCGCGCCGATGCGGCAATTT 803
QY 754 GATGGAGTTTAAAGAACTGGGAGGACAGGGGCA 794
Db 804 GACGACCTCTCCGGGGTGACACGGGCAAGACAGGAGAA 844

RESULT 4
US-08-446-875-15
; Sequence 15, Application US/08446875
; Patent No. 5858751
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, Sorge
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzhradszky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Identification and Synthesis of sialyltransferases
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.875
; FILING DATE: July 12, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/102,385
; FILING DATE: August 4, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 111-197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1020 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORGANISM: NO
US-08-446-875-15
Query Match 3.9%; Score 50.6; DB 2; Length 1020;
Best Local Similarity 53.9%; Pred. No. 3.2e-06;
Matches 104; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 389 TGGACTGTGACCTTTGTGCATAGTGTCAAACTCAGTCAGATGCTTGCCAGAAAGTGG 448
Db 410 TGGGCTCCGGCGCTGCGCGGTTGGGCACTGGGCACTGAGGAGCTTCCTATG 469
QY 449 GAATAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Db 470 GGCCTGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 529
QY 509 ATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
Db 530 TTGAAGTGTGTTGGGCAAGACCCACCACCTGCTGGTGTGATGATGATGATGATGATGAT 589
```

```
QY 569 TTTTGTCTAAAAA 581
Db 590 AGCTGCCACATAA 602

RESULT 5
US-08-446-875-1
; Sequence 1, Application US/08446875
; Patent No. 5858751
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, Sorge
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzhradszky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.875
; FILING DATE: July 12, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/102,385
; FILING DATE: August 4, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 111-197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: porcine
; TISSUE TYPE: liver, submaxillary glands
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 91..119
; OTHER INFORMATION: /product= "porcine Gal Beta 1,3
; OTHER INFORMATION: Galnac alpha 2,3 sialyltransferase"
US-08-446-875-1
Query Match 3.8%; Score 49.8; DB 2; Length 1218;
Best Local Similarity 55.5%; Pred. No. 6.6e-06;
Matches 96; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 393 CTGTGACCTTTGTGCATAGTGTCAAACTCAGTCAGATGCTTGCCAGAAAGTGGAAA 452
Db 513 CTGCGCGCGCTGCGCGCTGCTGGGCAACTGCGGCACTGAGAGATCTCTACTATGGCC 572
```

QY 453 TGAGATAGATCGATCCTCTGCAATTTGGAGATGAACAATGCCACCAAGAGTTATGA 512
DB 573 TCAGATAGACAGCCACGACTCTGCTGAGATGAACAAGCCCCCACCAGGAGGTTTGA 632
QY 513 AGAAGATGTCGGCGCCATGACCATGATTTCGAGTTGTGTCCCATACCAAGCGTTC 565
DB 633 GGCGAGCTCGGAGCAGACCAACCAATTCGTGTGATCCCGAGAGCTTCC 685

RESULT 6
US-08-102-385G-1
; Sequence 1, Application US/08102385G
; Patent No. 5962294
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, Sorge
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzhiradzky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, 38th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/102,385G
; FILING DATE: 04-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925369
; FILING DATE: 04-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 97-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310)788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: porcine
; TISSUE TYPE: liver, submaxillary glands
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 91..1119
; OTHER INFORMATION: /product= "porcine Gal Beta 1,3
; OTHER INFORMATION: Galnac alpha 2,3 sialyltransferase"
US-08-102-385G-1

Query Match 3.8%; Score 49.8; DB 2; Length 1218;
Best Local Similarity 55.5%; Pred. No. 6.6e-06;
Matches 96; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 393 CTGTGACCTTTGTGCCATAGTGTCAACTCAGGTGATGTTGGCCAGAGGTGGAA 452

DB 513 CUGCGGGCGTGGCCGTCGTCGGCACTCGGGCACTGAAGGAGTCTCTACTATGGGCC 572
QY 453 TGAGATAGATCGATCCTCTGCAATTTGGAGATGAACAATGCCACCAAGAGTTATGA 512
DB 573 TCAGATAGACAGCCACGACTCTGCTGAGATGAACAAGCCCCCACCAGGAGGTTTGA 632
QY 513 AGAAGATGTCGGCGCCATGACCATGATTTCGAGTTGTGTCCCATACCAAGCGTTC 565
DB 633 GGCGAGCTCGGAGCAGACCAACCAATTCGTGTGATCCCGAGAGCTTCC 685

RESULT 7
US-07-991-587A-1
; Sequence 1, Application US/07991587A
; Patent No. 5384249
; GENERAL INFORMATION:
; APPLICANT: Sasaki, Katsutoshi
; APPLICANT: Watanabe, Etsuyo
; APPLICANT: Nishi, Tatsunari
; APPLICANT: Sekine, Susumu
; APPLICANT: Hanai, No. 5384249uo
; APPLICANT: Hasegawa, Mamoru
; TITLE OF INVENTION: '2 3 Sialyltransferase
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitzpatrick, Cella, Harper & Scinto
; STREET: 277 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10172
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44mb
; COMPUTER: IBM PC
; OPERATING SYSTEM: Dos 3.3
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,587A
; FILING DATE: 19930526
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-333661/1991
; APPLICATION NUMBER: JP-091044/1992
; FILING DATE: 17-12-1991
; FILING DATE: 10-04-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence S. Perry
; REGISTRATION NUMBER: 31,865
; REFERENCE/DOCKET NUMBER: 1580.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-2400
; TELEFAX: 212-758-2982
; TELEX: 236262
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1919
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; CELL LINE: TYH cell
; CELL TYPE: histiocytoma cell
US-07-991-587A-1

Query Match 2.8%; Score 36.8; DB 1; Length 1919;
Best Local Similarity 51.2%; Pred. No. 0.18;
Matches 86; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 388 CTGGAAGCTGACCTTTGTGCCATAGTGTCAACTCAGGTGATGTTGGCCAGAGGTG 447


```

US-08-957-742-4
; Sequence 4, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAc ` 2,8-
; TITLE OF INVENTION: SIALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE: April 3, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/626,994
; FILING DATE: April 3, 1996
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE: April 3, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/626,994
; FILING DATE: April 3, 1996
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1048 bases
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-957-742-4

Query Match 2.8%; Score 36.6; DB 3; Length 1048;
Best Local Similarity 52.3%; Pred. No. 0.14;
Matches 81; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 381 TTGCAACTGGACTGTGACCTTTGGCCATAGTGCACAACTCAGGTCAGATGGTTGGCCA 440
Db 352 TATGAATAAGCGTTATATATGTTGTGCTGTGTTGGAACAGTSGAATCTTGACAGGGAG 411
QY 441 GAAGGTGGGAATGAGATAGATGATCCTCTCGCATTTGGAGATGAACAATGCCCCAC 500
Db 412 TCAGTGTGACAGAAATAGATGATGATTTGTTCTCGATGCAATTTGGCCCGAC 471
QY 501 CAAAGGTATGAAAGATGTCGCCCGCATGACCA 535
Db 472 AGAGGCTTCCCAAGATGTTGGAAGGAAACCA 506

RESULT 11
US-08-626-994A-2
; Sequence 2, Application US/08626994A
; Patent No. 5798244
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAc ` 2,8-
; TITLE OF INVENTION: SIALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place

```

```

; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/626,994A
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 bases
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-626-994A-2

Query Match 2.8%; Score 36.6; DB 1; Length 1660;
Best Local Similarity 52.3%; Pred. No. 0.19;
Matches 81; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 381 TTGCAACTGGACTGTGACCTTTGGCCATAGTGCACAACTCAGGTCAGATGGTTGGCCA 440
Db 536 TATGAATAAGCGTTATATATGTTGTGCTGTGTTGGAACAGTSGAATCTTGACAGGGAG 595
QY 441 GAAGGTGGGAATGAGATAGATGATCCTCTCGCATTTGGAGATGAACAATGCCCCAC 500
Db 596 TCAGTGTGACAGAAATAGATGATGATTTGTTCTCGATGCAATTTGGCCCGAC 655
QY 501 CAAAGGTATGAAAGATGTCGCCCGCATGACCA 535
Db 656 AGAGGCTTCCCAAGATGTTGGAAGGAAACCA 690

RESULT 12
US-08-957-742-2
; Sequence 2, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAc ` 2,8-
; TITLE OF INVENTION: SIALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE: April 3, 1995
; CLASSIFICATION: 435

```

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/626,994
/ FILING DATE: April 3, 1996
/ APPLICATION NUMBER: JP 77469/1995
/ FILING DATE: April 3, 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Arnold Turk
/ REGISTRATION NUMBER: 33,094
/ REFERENCE/DOCKET NUMBER: P14595
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 716-1191
/ TELEFAX: (703) 716-1180
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1660 bases
/ TYPE: nucleic acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: linear
/ US-08-957-742-2

Query Match          2.8%; Score 36.6; DB 3; Length 1660;
Best Local Similarity 52.3%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 74;

QY 381 TTTCGAACGTGGACTGTGACCTTTGTGCATAGTGTCAAACTCAGGTCAGATGGTTGGCCA 440
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 441 GAAGTGGGAATAGATAGATGATGCTCTCGCATTTGGAGATGAAATGCCGCCAC 500
Db      ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 596 TCAGTGTGACAGAAGATAGATAATATGATTTGTTCTCGATGCAATTTTCCCGCGAC 655
Db      ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 501 CAAAGGTTATGAAGAAGATGTCGCGCGCATGACCA 535
Db      ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 656 AGAGGCTTCCACAGAATGTTGGAAGGAACCA 690
Db      ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 13
US-09-996-243-346
; Sequence 346, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/0495787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
```

; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26

; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 2.8%; Score 35.6; DB 4; Length 2528;
Best Local Similarity 55.7%; Pred. No. 0.53;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 403 TGTGCATAGTGTCAAACTCAGGTCAGATGTTGGCCAGAGAGTGGGAAATGAGATAGAT 462
Db 1238 TGTGCGGTGGTGGCAACGGGGGCATCTCTGAACAACCTCCACATGGGCCAGAGATAGAC 1297

QY 463 CGATCCCTCTGCAATTTGGAGAAATGAACAATGCCCAACCAAGGTTATGAAGAAGATGTC 522
Db 1298 AGTCACGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAGGCTAGCAACAGGATGTG 1357

QY 523 GG 524
Db 1358 GG 1359

RESULT 14
US-09-539-333D-1
; Sequence 1: Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilva
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bibain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539.333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm

```

/ SEQ ID NO 1
/ LENGTH: 319608
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 31..1107
/ OTHER INFORMATION: 5'regulatory region g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 1108..1289
/ OTHER INFORMATION: exon A g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 14877..14920
/ OTHER INFORMATION: exon B g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 18778..18862
/ OTHER INFORMATION: exon Bbis g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 25593..25740
/ OTHER INFORMATION: exon C g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 29388..29502
/ OTHER INFORMATION: exon D g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 29967..30282
/ OTHER INFORMATION: exon E g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 64666..64812
/ OTHER INFORMATION: exon F g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 65505..65853
/ OTHER INFORMATION: exon G g35018 gene
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 65854..67854
/ OTHER INFORMATION: 3'regulatory region g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 94124..94964
/ OTHER INFORMATION: exon g35017
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 201188..201234
/ OTHER INFORMATION: exon S g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 214676..214793
/ OTHER INFORMATION: exon T g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 215702..215746
/ OTHER INFORMATION: exon U g35030 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216836..216915
/ OTHER INFORMATION: exon V g35030 gene
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 213818..215818
/ OTHER INFORMATION: 3'regulatory region g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 213819..215941
/ OTHER INFORMATION: exon R complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 215819..215975
/ OTHER INFORMATION: exon Rbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216661..216952
/ OTHER INFORMATION: exon Qbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 216661..217061
/ OTHER INFORMATION: exon Q complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 217027..217061
/ OTHER INFORMATION: exon Q1 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 229647..229742
/ OTHER INFORMATION: exon X complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 230408..230721
/ OTHER INFORMATION: exon P complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231272..231412
/ OTHER INFORMATION: exon Obis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231787..231880
/ OTHER INFORMATION: exon O2 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231870..231879
/ OTHER INFORMATION: exon O1 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 234174..234321
/ OTHER INFORMATION: exon O complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 237406..237428
/ OTHER INFORMATION: exon Nbis complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 239719..239807
/ OTHER INFORMATION: exon N2 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 239719..239853
/ OTHER INFORMATION: exon N complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240569
/ OTHER INFORMATION: exon M1117 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240596
/ OTHER INFORMATION: exon M1090 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240617
/ OTHER INFORMATION: exon M1069 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240644
/ OTHER INFORMATION: exon MS2 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 240528..240824
/ OTHER INFORMATION: exon M862 complement g34872 gene
/ FEATURE:
/ NAME/KEY: exon

```



```
/ LOCATION: 173358
/ OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 189957
/ OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 197163
/ OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 200778
/ OTHER INFORMATION: 8-303-235 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 202651
/ OTHER INFORMATION: 8-300-221 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 202679
/ OTHER INFORMATION: 8-300-193 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 203378
/ OTHER INFORMATION: 8-299-128 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 204138
/ OTHER INFORMATION: 8-296-213 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 204605
/ OTHER INFORMATION: 8-252-190 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 204934
/ OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 205206
/ OTHER INFORMATION: 8-295-248 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 205329
/ OTHER INFORMATION: 8-295-125 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 206064
/ OTHER INFORMATION: 8-293-130 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 206545
/ OTHER INFORMATION: 8-292-198 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 207313
/ OTHER INFORMATION: 8-251-322 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 208285
/ OTHER INFORMATION: 8-289-322 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 208960
/ OTHER INFORMATION: 8-287-249 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 209123
/ OTHER INFORMATION: 8-287-86 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 209631
/ OTHER INFORMATION: 8-285-319 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 210361
/ OTHER INFORMATION: 8-283-278 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 210463
/ OTHER INFORMATION: 8-283-176 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 210486
/ OTHER INFORMATION: 8-283-153 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 210583
/ OTHER INFORMATION: 8-283-56 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 210879
/ OTHER INFORMATION: 8-282-345 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 210964
```

```
/ OTHER INFORMATION: 8-282-260 : polymorphic base G or T
/ NAME/KEY: allele
/ LOCATION: 210979
/ OTHER INFORMATION: 8-282-245 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 211050
/ OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAGGAGGAAGAAGA
/ NAME/KEY: allele
/ LOCATION: 211132
/ OTHER INFORMATION: 8-282-92 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 211247
/ OTHER INFORMATION: 8-281-367 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 211315
/ OTHER INFORMATION: 8-281-299 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 211366
/ OTHER INFORMATION: 8-281-248 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 212520
/ OTHER INFORMATION: 8-279-197 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 212821
/ OTHER INFORMATION: 8-278-289 : polymorphic base C or T
/ NAME/KEY: allele

Query Match 2.7%; Score 35.4; DB 4; Length 319608;
Best Local Similarity 49.7%; Pred. No. 16;
Matches 90; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1027 TTCTATTCTGAGGAGCAGGGTCTGTGTGAATTGCACACAGGAGGACCAATCCCTT 1086
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116729 TTTTCATCCGTGAGGGCCCTCAGGCTATATACAAACAATACACAAAACCTGCATCAT 116788
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1087 GCCCTGATACAGGCAACCTGAGTGTCTTCTCTGCTCAGAACTAGTGTGACTA 1146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116789 TACCTGGGACATGTTATGGGCTACTAAGTCGATCCTTTTAGAAGAGCAGTGAATG 116848
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1147 TGTGGCCTACCTCACATGTTTGTGTACACCTACACAGGAAAAAGAAATGTCCTTT 1206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116849 GATTGTTTCTCTGACAAATTTGTAGTACAGTAAAAATAAATGGGAAAAATTTT 116908
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1207 T 1207
Db 116909 T 116909
```

Search completed: September 12, 2003, 15:04:52
Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 12:31:53 ; Search time 416 Seconds
(without alignments)
8396.810 Million cell updates/sec

Title: US-09-714-936-218

Perfect score: 1294

Sequence: 1 ccggaattccgggtcgacg.....ttttctctctttttttt 1294

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1294	100.0	1294	24	ABQ99471 Human coding seque
2	743.4	57.4	1106	22	ACA56192 Human DNA encoding
3	743.4	57.4	1106	25	ACA57950 Human PRO1717 cDNA
4	743.4	57.4	1106	25	ABX98420 Human cDNA encodin
5	743.4	57.4	1106	25	ABX98922 Novel human secret
6	743.4	57.4	1106	25	ACA05967 Human secreted/tr
7	743.4	57.4	1106	25	ABX98011 Human PRO polynucl
8	743.4	57.4	1106	25	ABX78795 Human PRO polynucl

9	743.4	57.4	1106	25	ABX75808 Human cDNA encodin
10	743.4	57.4	1106	25	ABX77013 Human PRO polynucl
11	743.4	57.4	1106	25	ABX16853 Human cDNA encodin
12	743.4	57.4	1122	22	AAO09952 Human drug metabol
13	714	55.2	2575	24	AAO34123 Human secreted pro
14	713.8	55.2	3189	24	AAO34106 Human secreted pro
15	652.4	50.4	696	21	AAO76478 Human ORFX ORF2033
16	608	47.0	725	22	AAO34124 Human secreted pro
17	554	42.8	672	22	AAO1230 cDNA encoding nove
18	258	19.9	683	22	AAO20502 Human breast cance
19	190.4	14.7	209	23	ABV59345 Human prostate exp
20	177.2	13.7	646	22	AAO11604 Human breast cance
21	157.4	12.2	211	23	AAV59361 Human prostate exp
22	98	7.6	906	22	AAO28685 Human protein HP03
23	98	7.6	1416	21	AAO65393 Human carbohydrate
24	98	7.6	1556	21	AAO87726 Human secreted pro
25	98	7.6	1556	22	AAO64008 cDNA encoding huma
26	98	7.6	1687	22	AAO28695 Human protein HP03
27	98	7.6	1735	21	AAO78385 Human secreted pro
28	88.4	6.8	1124	22	AAO53252 Human polynucleoti
29	88.4	6.8	2048	24	ABO78463 cDNA encoding huma
30	88.4	6.8	2056	22	AAO18037 Human sialyltransf
31	88.4	6.8	2075	24	AAO35226 Human TRNFR-11 cDN
32	87.4	6.8	549	23	ABV60266 Human prostate exp
33	83.4	6.4	814	22	ABA21356 Human nervous syst
34	83.4	6.4	814	22	AAO78174 Human immune/haema
35	83	6.4	1091	22	AAO16038 Human polynucleoti
36	82.6	6.4	2350	22	AAO16695 Human cDNA sequenc
37	82.6	6.4	2362	21	AAO76092 Human ORFX ORF1647
38	82.6	6.4	2401	21	AAO37038 Human PRO1359 (UNQ
39	82.6	6.4	2401	22	AAO46027 Human DNA encoding
40	82.6	6.4	2401	22	AAO54240 DNA encoding prote
41	82.6	6.4	2401	25	AAO57785 Human cDNA encodin
42	82.6	6.4	2401	25	ABX98255 Human cDNA encodin
43	82.6	6.4	2401	25	ABX98757 Novel human secret
44	82.6	6.4	2401	25	ACA05802 Human secreted/tr
45	82.6	6.4	2401	25	ABX97846 Human PRO polynucl

ALIGNMENTS

RESULT 1
ID ABQ99471 standard; cDNA; 1294 BP.

XX AC ABQ99471;
XX DT 25-FEB-2003 (first entry)
XX DE Human coding sequence SEQ ID 204.

XX KW Human; expressed sequence tag; EST;
KW haematopoietic disorder; central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
KW antileukemic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytoskeletal; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200259260-A2.

XX PD 01-AUG-2002.

XX PF 16-NOV-2001; 2001WO-US42950.

XX PR 17-NOV-2000; 2000US-0714936.

XX PA (HYSE-) HYSEQ INC.

XX XX

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Tang Y, Mehrman T, Drmanac RT;
 XX WPI; 2002-590824/63.
 DR N-PSDB; ABP64885.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 XX
 PS Claim 1; SEQ ID 204; 394pp; English.
 XX
 CC The present invention relates to novel human coding sequences
 CC (AB099268-AB099608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1294 Bp; 320 A; 312 C; 313 G; 349 T; 0 other;

Query Match 100.0%; Score 1294; DB 24; Length 1294;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGGAATTCCGGGTGCGAGATTCGTCGCGGTCCCTTATTGGATCTCGGGAATG 60
 DB 1 CCGGAATTCCGGGTGCGAGATTCGTCGCGGTCCCTTATTGGATCTCGGGAATG 60
 QY 61 GGGCTGGAGAGTCTCGCGGTACACAGCCTCAGCCTGCCCCCAGGACATGCCCCGTGAC 120
 DB 61 GGGCTGGAGAGTCTCGCGGTACACAGCCTCAGCCTGCCCCCAGGACATGCCCCGTGAC 120
 QY 121 CCAGCGCGCCGCTCGGTGGCAGAGGGCGCGGAGCGCATGGCTGCTGATCCTG 180
 DB 121 CCAGCGCGCCGCTCGGTGGCAGAGGGCGCGGAGCGCATGGCTGCTGATCCTG 180
 QY 181 AAGAGAAAGTCTGTGATCTGTGAGCTTCATACAGCGCTTCCTTTCTGCTGGTTGTG 240
 DB 181 AAGAGAAAGTCTGTGATCTGTGAGCTTCATACAGCGCTTCCTTTCTGCTGGTTGTG 240
 QY 241 CTTCTTGAATGAAGTGAATTTCCATTGCTACTAACTGCTTTGGACACCTGGTACA 300
 DB 241 CTTCTTGAATGAAGTGAATTTCCATTGCTACTAACTGCTTTGGACACCTGGTACA 300
 QY 301 AAGTGGATACCATCTCTTACACATACAGGGCGCCCTCGCACTCACTATGATACATA 360
 DB 301 AAGTGGATACCATCTCTTACACATACAGGGCGCCCTCTGCACTCACTATGATACATA 360
 QY 361 AATGTGAGACACAGAGCCTTTGCACTGGACTGTGACCTTTGTGCCATAGTGTCAAC 420
 DB 361 AATGTGAGACACAGAGCCTTTGCACTGGACTGTGACCTTTGTGCCATAGTGTCAAC 420

QY 421 TCAGGTGATGTTGGCCAGAGGTGGGAAATGAGATAGATCGATCCTCTCTCATTTGG 480
 DB 421 TCAGGTGATGTTGGCCAGAGGTGGGAAATGAGATAGATCGATCCTCTCTCATTTGG 480
 QY 481 AGAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTGGCCCATGACCATGATT 540
 DB 481 AGAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTGGCCCATGACCATGATT 540
 QY 541 CGAGTGTGTCCTCCATPACAGCGTTCCCTCTTTGCTTAAAAAACCCCTGATTTTTCAG 600
 DB 541 CGAGTGTGTCCTCCATPACAGCGTTCCCTCTTTGCTTAAAAAACCCCTGATTTTTCAG 600
 QY 601 GAAGCGAATACACTATTTGTTATTTGGGACCTTTCCGCAATATGAGAAAGATGCG 660
 DB 601 GAAGCGAATACACTATTTGTTATTTGGGACCTTTCCGCAATATGAGAAAGATGCG 660
 QY 661 AATGGCATCGTTTACAACATGTTGAAAAAGACAGTTGGTATCTATCCGAATGCCAAATA 720
 DB 661 AATGGCATCGTTTACAACATGTTGAAAAAGACAGTTGGTATCTATCCGAATGCCAAATA 720
 QY 721 TAGTGACACACAGAGAACCGCATGAGTTACTGTGATGAGTTTAAAGAGAACTGG 780
 DB 721 TAGTGACACACAGAGAACCGCATGAGTTACTGTGATGAGTTTAAAGAGAACTGG 780
 QY 781 AAGGACAGGGGCGCATGCAAGCGGACTGCTGATTTCTACAGACACTTTTAAAGCATTAC 840
 DB 781 AAGGACAGGGGCGCATGCAAGCGGACTGCTGATTTCTACAGACACTTTTAAAGCATTAC 840
 QY 841 AGTGTGGAAGTGAAGTAACTTTCGGTCCCTCTTACAGCATCCAGTCCACTTGTGTA 900
 DB 841 AGTGTGGAAGTGAAGTAACTTTCGGTCCCTCTTACAGCATCCAGTCCACTTGTGTA 900
 QY 901 ATGTCATAGCGATATAAATCTGTCAGCGGCGAGGATTCATGTCATCTCTGCTCTCTCC 960
 DB 901 ATGTCATAGCGATATAAATCTGTCAGCGGCGAGGATTCATGTCATCTCTGCTCTCTCC 960
 QY 961 TTCACGTAAACAAATCTCATTTGTTGATGGCATATGGCAGCAAGCATCCCAACACCCAGA 1020
 DB 961 TTCACGTAAACAAATCTCATTTGTTGATGGCATATGGCAGCAAGCATCCCAACACCCAGA 1020
 QY 1021 GTGGTGTCTTATTTCTGAGGAGCAGGCTCTGTGTGAATTCACACACAGGAGCA 1080
 DB 1021 GTGGTGTCTTATTTCTGAGGAGCAGGCTCTGTGTGAATTCACACACAGGAGCA 1080
 QY 1081 TCCCTGCTGCTGATACAGCAACCTGAGTCTTGTAGTCTCTCTGCTCTGCTGCTGCTGCT 1140
 DB 1081 TCCCTGCTGCTGATACAGCAACCTGAGTCTTGTAGTCTCTCTGCTCTGCTGCTGCTGCT 1140
 QY 1141 TGACTATGCGCTTACCTCAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
 DB 1141 TGACTATGCGCTTACCTCAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
 QY 1201 TCCCTTTGATTCATGCTGTTGATAGATGTTTCATCCAAATTTGAATGACATGTGACCAAG 1260
 DB 1201 TCCCTTTGATTCATGCTGTTGATAGATGTTTCATCCAAATTTGAATGACATGTGACCAAG 1260
 QY 1261 TAGTGTCTTCCCT 1294
 DB 1261 TAGTGTCTTCCCT 1294

RESULT 2
 AAS46192
 ID AAS46192 standard; cDNA; 1106 BP.
 XX
 XX AAS46192;
 AC
 XX
 DT 18-DEC-2001 (first entry)
 XX
 XX Human DNA encoding PRO polypeptide sequence #268.
 DE
 XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;

KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.

XX Homo sapiens.

OS W0200168848-A2.

PN 20-SEP-2001.

PD 28-FEB-2001; 2001WO-US06520.

PF 01-MAR-2000; 2000WO-US05601.

PR 02-MAR-2000; 2000WO-US05841.

PR 03-MAR-2000; 2000US-187202P.

PR 06-MAR-2000; 2000US-186968P.

PR 14-MAR-2000; 2000US-189320P.

PR 14-MAR-2000; 2000US-189328P.

PR 15-MAR-2000; 2000WO-US06884.

PR 21-MAR-2000; 2000US-190828P.

PR 21-MAR-2000; 2000US-191007P.

PR 21-MAR-2000; 2000US-191048P.

PR 21-MAR-2000; 2000US-191314P.

PR 28-MAR-2000; 2000US-192655P.

PR 29-MAR-2000; 2000US-193032P.

PR 29-MAR-2000; 2000US-193053P.

PR 30-MAR-2000; 2000WO-US08439.

PR 04-APR-2000; 2000US-194449P.

PR 04-APR-2000; 2000US-194647P.

PR 11-APR-2000; 2000US-195975P.

PR 11-APR-2000; 2000US-196000P.

PR 11-APR-2000; 2000US-196187P.

PR 11-APR-2000; 2000US-196690P.

PR 18-APR-2000; 2000US-196820P.

PR 18-APR-2000; 2000US-198121P.

PR 25-APR-2000; 2000US-198585P.

PR 25-APR-2000; 2000US-199397P.

PR 25-APR-2000; 2000US-199550P.

PR 03-MAY-2000; 2000US-201516P.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.

PR 05-JUN-2000; 2000US-209832P.

PR 28-JUL-2000; 2000WO-US20710.

PR 22-AUG-2000; 2000US-084848P.

PR 24-AUG-2000; 2000WO-US23328.

PR 08-NOV-2000; 2000WO-US30952.

PR 01-DEC-2000; 2000WO-US32678.

PR 20-DEC-2000; 2000WO-US34956.

XX (GETH) GENENTECH INC.

PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2001-602746/68.

XX P-PSDB; ANU29291.

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -

XX Claim 2; Fig 535; 774pp; English.

XX Sequences AAS45925-AA846231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
XX invention can be used to detect the presence of a tumour in a mammal by
XX comparing the level of expression of a PRO polypeptide in a test sample
XX of cells from the animal and a control sample of normal cells, whereby a
XX higher level of expression in the test sample indicates the presence of a
XX tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,

CC pigs, goats and rabbits but are preferably human. The polypeptides can be
CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 1106 BP; 325 A; 226 C; 260 G; 295 T; 0 other;

Query Match 57.4%; Score 743.4; DB 22; Length 1106;

Best Local Similarity 99.9%; Pred. No. 4.6e-233;

Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTGCGGGAATGTGGGCTGGAGAGTCTGCGGTGCTACAGCCTCCAGCCTGCC 103

Db 1 TGGATCTGCGGGAATGTGGGCTGGAGAGTCTGCGGTGCTACAGCCTCCAGCCTGCC 60

QY 104 CCAGGACTGCCCTGACCCAGCCAGCGGCCGCTGCTCGTGGCAGGAGCGCGGAGCG 163

Db 61 CCAGGACTGCCCTGACCCAGCCAGCGGCCGCTGCTCGTGGCAGGAGCGCGGAGCG 120

QY 164 CCATGGCTGCATCCTGAGAGGAAAGTCTGTGATTCCTGAGCTTCATAGACGCTTC 180

Db 121 CCATGGCTGCATCCTGAGAGGAAAGTCTGTGATTCCTGAGCTTCATAGACGCTTC 180

QY 224 TTTTCTCTGCTGGTGTGGCTCTTGTAAATGAAGTGAATTCCTCATTTGCTACTAACTGCT 283

Db 181 TTTTCTCTGCTGGTGTGGCTCTTGTAAATGAAGTGAATTCCTCATTTGCTACTAACTGCT 240

QY 284 TTGACAACTGTCACAAAGTGGATACATTCCTTACACATACAGGCGCCCTTCGAA 343

Db 241 TTGACAACTGTCACAAAGTGGATACATTCCTTACACATACAGGCGCCCTTCGAA 300

QY 344 CTCACATATGATACATAATATGAGACACAAAGAGCTTTGCAACTGGACGTGACCTTT 403

Db 301 CTCACATATGATACATAATATGAGACACAAAGAGCTTTGCAACTGGACGTGACCTTT 360

QY 404 GTGCCATAGTGTCAAACACTCAGTCAGATGGTTGCCAGAAAGTGGGAAATGAGATAGTC 463

Db 361 GTGCCATAGTGTCAAACACTCAGTCAGATGGTTGCCAGAAAGTGGGAAATGAGATAGTC 420

QY 464 GATCCTCTCTCATTTGGAGATGAACAATGCCGCCACCAAGGTTATGAAGAAGATGTC 523

Db 421 GATCCTCTCTCATTTGGAGATGAACAATGCCGCCACCAAGGTTATGAAGAAGATGTC 480

QY 524 GCCGCATGACCATGATTCGAGTTGTGTCCTCATCCAGCGTTCCTCTTTGCTAAAAAACC 583

Db 481 GCCGCATGACCATGATTCGAGTTGTGTCCTCATCCAGCGTTCCTCTTTGCTAAAAAACC 540

QY 584 CTGATTATTTTCAAGGAGGGAATACACTACTATTTGTTGTTATTTGGGACCTTTCGCA 643

Db 541 CTGATTATTTTCAAGGAGGGAATACACTACTATTTGTTGTTATTTGGGACCTTTCGCA 600

QY 644 ATATGAGGAAAGATGGCAATGTCATCGTTTACACATGTTGAAAGACAGCTTGTATCT 703

Db 601 ATATGAGGAAAGATGGCAATGTCATCGTTTACACATGTTGAAAGACAGCTTGTATCT 660

QY 704 ATCCGAATGCCCAATPATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 763

Db 661 ATCCGAATGCCCAATPATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 720

QY 764 TTAAGAAGGAACTGGGAAGGACAG 788

Db 721 TTAAGAAGGAACTGGGAAGGACAG 745

RESULT 3

ACA57950

ID ACA57950 standard; cDNA; 1106 BP.

XX

AC ACA57950;
XX 10-JUN-2003 (first entry)
XX Human PRO7177 cDNA.
XX
XX Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood; gene;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy; ss.
XX
OS Homo sapiens.
XX US2003036143-A1.
PN 20-FEB-2003.
XX
XX 02-JUL-2002; 2002US-0187600.
XX 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US23108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263B.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 13-OCT-1997; 97US-064103P.
PR 13-OCT-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-070024P.
PR 22-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 18-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 01-APR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 18-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.

PR	22-JUN-1998;	98US-090254P.	
PR	24-JUN-1998;	98US-090429P.	
PR	24-JUN-1998;	98US-090435P.	
PR	24-JUN-1998;	98US-090444P.	
PR	24-JUN-1998;	98US-090461P.	
PR	24-JUN-1998;	98US-090533P.	
PR	24-JUN-1998;	98US-090540P.	
PR	25-JUN-1998;	98US-090676P.	
PR	25-JUN-1998;	98US-090678P.	
PR	25-JUN-1998;	98US-090688P.	
PR	25-JUN-1998;	98US-090690P.	
PR	25-JUN-1998;	98US-090694P.	
PR	25-JUN-1998;	98US-090695P.	
PR	25-JUN-1998;	98US-090696P.	
PR	26-JUN-1998;	98US-090862P.	
PR	26-JUN-1998;	98US-090863P.	
PR	26-JUN-1998;	98US-091010P.	
PR	01-JUL-1998;	98US-091359P.	
PR	01-JUL-1998;	98US-091344P.	
PR	02-JUL-1998;	98US-091478P.	
PR	02-JUL-1998;	98US-091486P.	
PR	02-JUL-1998;	98US-091626P.	
PR	02-JUL-1998;	98US-091628P.	
PR	02-JUL-1998;	98US-091632P.	
PR	24-JUL-1998;	98US-094006P.	
PR	04-AUG-1998;	98US-095282P.	
PR	10-AUG-1998;	98US-095998P.	
PR	10-AUG-1998;	98US-096012P.	
PR	17-AUG-1998;	98US-096757P.	
PR	17-AUG-1998;	98US-096766P.	
PR	17-AUG-1998;	98US-096867P.	
PR	17-AUG-1998;	98US-096891P.	
PR	17-AUG-1998;	98US-096897P.	
PR	18-AUG-1998;	98US-096949P.	
PR	18-AUG-1998;	98US-096959P.	
PR	18-AUG-1998;	98US-097022P.	
PR	26-AUG-1998;	98US-097952P.	
PR	26-AUG-1998;	98US-097954P.	
PR	26-AUG-1998;	98US-097955P.	
PR	26-AUG-1998;	98US-097971P.	
PR	26-AUG-1998;	98US-097974P.	
PR	26-AUG-1998;	98US-098014P.	
PR	01-SEP-1998;	98US-098716P.	
PR	01-SEP-1998;	98US-098723P.	
PR	02-SEP-1998;	98US-098803P.	
PR	02-SEP-1998;	98US-098821P.	
PR	02-SEP-1998;	98US-098843P.	
PR	09-SEP-1998;	98US-099602P.	
PR	10-SEP-1998;	98US-099741P.	
PR	10-SEP-1998;	98US-099754P.	
PR	10-SEP-1998;	98US-099763P.	
PR	10-SEP-1998;	98US-099812P.	

Query Match 57.4%; Score 743.4; DB 25; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.6e-233;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	44	TGGATCTCGCGGAACTGGCTGGAGAGTCTCTGGCTGGTACACGCTCCAGCTGCC 103
Db	1	TGGATCTCGCGGAACTGGCTGGAGAGTCTCTGGCTGGTACACGCTCCAGCTGCC 60
QY	104	CCAGGACTGCCCTGACCCAGCGCGCTGGTGGAGAGGCGCGCGGCGGAGGG 163
Db	61	CCAGGACTGCCCTGACCCAGCGCGCTGGTGGAGAGGCGCGCGGCGGAGGG 120
QY	164	CCATGGCTGCATCTCGAAGAAAGTCTGTGATCTGTGAGCTTCATAGCAGCGTTC 223
Db	121	CCATGGCTGCATCTCGAAGAAAGTCTGTGATCTGTGAGCTTCATAGCAGCGTTC 180
QY	224	TTTTCTCTGCTGGTGGCTCTTGAATGAAGTGAATTCCTACTACTAACTGCT 283
Db	181	TTTTCTCTGCTGGTGGCTCTTGTGAATGAAGTGAATTCCTACTACTAACTGCT 240

QY	284	TTGGCAACACCTGGTACAAAAGTGATACCAATCTCTCTACATACAGCGGCCCTTCGAA 343
Db	241	TTGGCAACACCTGGTACAAAAGTGATACCAATCTCTCTACATACAGCGGCCCTTCGAA 300
QY	344	CTCACTATGGATACATAAATGTGAAGACACAGAGCCTTTGCAACTGGACTGGACCTTT 403
Db	301	CTCACTATGGATACATAAATGTGAAGACACAGAGCCTTTGCAACTGGACTGGACCTTT 360
QY	404	GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAAATGAGATAGATC 463
Db	361	GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAAATGAGATAGATC 420
QY	464	GATCCTCTCTGATTTGGAGATGAACAATGCCGCCACCAAGGTTATGAGAGAGTGTGC 523
Db	421	GATCCTCTCTGATTTGGAGATGAACAATGCCGCCACCAAGGTTATGAGAGAGTGTGC 480
QY	524	GCGCATGACCATGATTCGAGTTGTCTCCATACACAGCGTTCCTCTTTGCTAAAAAAC 583
Db	481	GCGCATGACCATGATTCGAGTTGTCTCCATACACAGCGTTCCTCTTTGCTAAAAAAC 540
QY	584	CTGATTATTTTCAAGAGAGCGAATACACTACTATTATTTGGTTATTTGGGACCTTCCGCA 643
Db	541	CTGATTATTTTCAAGAGAGCGAATACACTACTATTATTTGGTTATTTGGGACCTTCCGCA 600
QY	644	ATATGAGGAAAGATGCGCAATGGCATCGTTTACACATGTTGAAAAAGACAGTTGTATCT 703
Db	601	ATATGAGGAAAGATGCGCAATGGCATCGTTTACACATGTTGAAAAAGACAGTTGTATCT 660
QY	704	ATCCGAATGCCCAATATACGTGACCACAGAGAGCGCATGAGTTACTGTGTGAGGTTT 763
Db	661	ATCCGAATGCCCAATATACGTGACCACAGAGAGCGCATGAGTTACTGTGTGAGGTTT 720
QY	764	TTAAGAGGAACATGCGGAAGGACAG 788
Db	721	TTAAGAGGAACATGCGGAAGGACAG 745
RESULT 4		
ABX98420		
ID	ABX98420	standard; cDNA; 1106 BP.
XX	AC	ABX98420;
XX	DT	19-MAY-2003 (first entry)
XX	DE	Human cDNA encoding a secreted/transmembrane protein, SEQ ID 535.
KW	Human; ss; gene; PRO; secreted protein; transmembrane protein;	
KW	cytostatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;	
KW	colon tumour; breast tumour; prostate tumour; rectal tumour;	
KW	cervical tumour; liver tumour; TNF-alpha release; arthritis;	
KW	tumour necrosis factor alpha; chondrocyte cell; bone disorder;	
KW	cartilage disorder; sports injury.	
XX	Homo sapiens.	
OS	XX	
PN	XX	US2003036156-A1.
XX	XX	20-FEB-2003.
PD	XX	02-JUL-2002; 2002US-0188767.
PF	XX	16-SEP-1998; 98WO-US19330.
XX	XX	07-OCT-1998; 98WO-US21141.
PR	XX	01-DEC-1998; 98WO-US25108.
PR	XX	08-MAR-1999; 99WO-US05028.
PR	XX	14-MAY-1999; 99WO-US10733.
PR	XX	02-JUN-1999; 99WO-US12252.
PR	XX	01-SEP-1999; 99WO-US20111.
PR	XX	15-SEP-1999; 99WO-US21090.
PR	XX	01-DEC-1999; 99WO-US28301.
PR	XX	02-DEC-1999; 99WO-US28551.
PR	XX	30-DEC-1999; 99WO-US31274.

PR 05-JAN-2000; 2000WO-US002119.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078866P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 31-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 01-APR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082566P.
PR 22-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 05-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088732P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090438P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.

PR	17-AUG-1998;	98US-096766P.	
PR	17-AUG-1998;	98US-096867P.	
PR	17-AUG-1998;	98US-096869P.	
PR	17-AUG-1998;	98US-096897P.	
PR	18-AUG-1998;	98US-096949P.	
PR	18-AUG-1998;	98US-096959P.	
PR	18-AUG-1998;	98US-097022P.	
PR	18-AUG-1998;	98US-097052P.	
PR	26-AUG-1998;	98US-097954P.	
PR	26-AUG-1998;	98US-097953P.	
PR	26-AUG-1998;	98US-097971P.	
PR	26-AUG-1998;	98US-097974P.	
PR	26-AUG-1998;	98US-098014P.	
PR	01-SEP-1998;	98US-098716P.	
PR	01-SEP-1998;	98US-098723P.	
PR	02-SEP-1998;	98US-098803P.	
PR	02-SEP-1998;	98US-098821P.	
PR	02-SEP-1998;	98US-098843P.	
PR	09-SEP-1998;	98US-099602P.	
PR	10-SEP-1998;	98US-099741P.	
Query Match 57.4%; Score 743.4; DB 25; Length 1106;			
Best Local Similarity 99.9%; Pred. No. 4.6e-233;			
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	44	TGGATCTGGGGAATGTGGGTGGAGAGTCTCTGCGTGTACAGGCTCCAGCCTTGCC	103
DB	1	TGGATCTGGGGAATGTGGGTGGAGAGTCTCTGCGTGTACAGGCTCCAGCCTTGCC	60
QY	104	CCAGGACTGCCCTGACCCAGGCGCGCCGCTGCTCGTGGCAGAGGGCCGCGAGGG	163
DB	61	CCAGGACTGCCCTGACCCAGGCGCGCCGCTGCTCGTGGCAGAGGGCCGCGAGGG	120
QY	164	CCATGCCCTGCATCTGAGAGAAAGTCTGTGATTCCTGCTACTACTAACTGCT	223
DB	121	CCATGCCCTGCATCTGAGAGAAAGTCTGTGATTCCTGCTACTACTAACTGCT	180
QY	224	TTTTCTGCTGTTGCTGCTGTTGTAATGAAGTGAATTCCTGCTACTACTAACTGCT	283
DB	181	TTTTCTGCTGTTGCTGCTGTTGTAATGAAGTGAATTCCTGCTACTACTAACTGCT	240
QY	284	TTGGACAACCTGGTACAAAGTGGATACCATCTTCCTACACATACAGCGGCCCTTCGAA	343
DB	241	TTGGACAACCTGGTACAAAGTGGATACCATCTTCCTACACATACAGCGGCCCTTCGAA	300
QY	344	CTCAGTATGGATACATTAATGTGAACACACAGAGCCTTGCAACTGGATGTGACCTTT	403
DB	301	CTCAGTATGGATACATTAATGTGAACACACAGAGCCTTGCAACTGGATGTGACCTTT	360
QY	404	GTGCCATAGTGTCAAACTCAGGTGATGTTGGCCAGAGGTGGGAAATGAGATAGATC	463
DB	361	GTGCCATAGTGTCAAACTCAGGTGATGTTGGCCAGAGGTGGGAAATGAGATAGATC	420
QY	464	GATCCTCTGCTGATTTGGAGATGAACAAATGCCCCCAAGGTTATGAAGAAGATGCG	523
DB	421	GATCCTCTGCTGATTTGGAGATGAACAAATGCCCCCAAGGTTATGAAGAAGATGCG	480
QY	524	GCCGATGACCATGATTCGAGTTGTGTCCTACACAGGCTTCCTCTTTGCTAAATAAC	583
DB	481	GCCGATGACCATGATTCGAGTTGTGTCCTACACAGGCTTCCTCTTTGCTAAATAAC	540
QY	584	CTGATTATTTTCAAGGAAGCAATACATCTATTGTTGTTATTGGGGACCTTTCCGCA	643
DB	541	CTGATTATTTTCAAGGAAGCAATACATCTATTGTTGTTATTGGGGACCTTTCCGCA	600
QY	644	ATATGAGGAAGATGGCAATGGATCTTTTACACATGTTGAAAAGACAGTGGTATCT	703
DB	601	ATATGAGGAAGATGGCAATGGATCTTTTACACATGTTGAAAAGACAGTGGTATCT	660
QY	704	ATCCGATGCCAAATATACGTGACACACAGAGAGCGCATGAGTACTGTGATGGAGTTT	763
DB	661	ATCCGATGCCAAATATACGTGACACACAGAGAGCGCATGAGTACTGTGATGGAGTTT	720

QY	764	TTAAGAGCAACTGGGAAGGACAG	788
DB	721	TTAAGAGCAACTGGGAAGGACAG	745
RESULT 5			
ABX98922			
ID	ABX98922	standard; cDNA; 1106 BP.	
XX	XX	ABX98922;	
XX	XX	20-MAY-2003 (first entry)	
DE	DE	Novel human secreted and transmembrane protein PRO7177 cDNA.	
KW	KW	Human; secreted protein; transmembrane protein; cytostatic;	
KW	KW	gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;	
KW	KW	adrenal tumour; lung tumour; colon tumour; breast tumour;	
KW	KW	prostate tumour; rectal tumour; cervical tumour; liver tumour;	
KW	KW	gene; ss.	
OS	OS	Homo sapiens.	
XX	XX	US2003036157-A1.	
XX	XX	20-FEB-2003.	
PF	PF	02-JUL-2002; 2002US-0188769.	
XX	XX	16-SEP-1998; 98WO-US19330.	
PR	PR	07-OCT-1998; 98WO-US21141.	
PR	PR	01-DEC-1998; 98WO-US25108.	
PR	PR	08-MAR-1999; 98WO-US05028.	
PR	PR	14-MAY-1999; 98WO-US10733.	
PR	PR	02-JUN-1999; 98WO-US12252.	
PR	PR	01-SEP-1999; 98WO-US20111.	
PR	PR	15-SEP-1999; 98WO-US21090.	
PR	PR	01-DEC-1999; 98WO-US28301.	
PR	PR	02-DEC-1999; 98WO-US28551.	
PR	PR	30-DEC-1999; 98WO-US31274.	
PR	PR	05-JAN-2000; 2000WO-US00219.	
PR	PR	18-FEB-2000; 2000WO-US04341.	
PR	PR	22-FEB-2000; 2000WO-US04342.	
PR	PR	24-FEB-2000; 2000WO-US04414.	
PR	PR	01-MAR-2000; 2000WO-US05004.	
PR	PR	02-MAR-2000; 2000WO-US05601.	
PR	PR	15-MAR-2000; 2000WO-US05841.	
PR	PR	30-MAR-2000; 2000WO-US06884.	
PR	PR	17-MAY-2000; 2000WO-US13705.	
PR	PR	22-MAY-2000; 2000WO-US14042.	
PR	PR	02-JUN-2000; 2000WO-US15264.	
PR	PR	28-JUL-2000; 2000WO-US20710.	
PR	PR	24-AUG-2000; 2000WO-US23328.	
PR	PR	08-NOV-2000; 2000WO-US30952.	
PR	PR	01-DEC-2000; 2000WO-US32678.	
PR	PR	20-DEC-2000; 2000WO-US34956.	
PR	PR	28-FEB-2001; 2001WO-US06520.	
PR	PR	01-JUN-2001; 2001WO-US17800.	
PR	PR	29-JUN-2001; 2001WO-US19692.	
PR	PR	09-JUL-2001; 2001WO-US21735.	
PR	PR	29-AUG-2001; 2001WO-US27099.	
PR	PR	18-SEP-1997; 97US-059263P.	
PR	PR	18-SEP-1997; 97US-059266P.	
PR	PR	17-OCT-1997; 97US-062250P.	
PR	PR	21-OCT-1997; 97US-063486P.	
PR	PR	24-OCT-1997; 97US-063120P.	
PR	PR	24-OCT-1997; 97US-063121P.	
PR	PR	28-OCT-1997; 97US-063540P.	
PR	PR	28-OCT-1997; 97US-063541P.	
PR	PR	28-OCT-1997; 97US-063544P.	

PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078866P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 31-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 21-APR-1998; 98US-082569P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087603P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088653P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.

PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.

Query Match 57.4%; Score 743.4; DB 25; Length 1106;

Best Local Similarity 99.9%; Pred. No. 4.6e-233;

Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 44 TGGATCTCGGGATGTGGCTGGAGAGTCTCTCCGTGTGTACCGCTCCAGCTCCGCC 103

Db 1 TGGATCTCGGGATGTGGCTGGAGAGTCTCTCCGTGTGTACCGCTCCAGCTCCGCC 60

Oy 104 CCAGGACTGCCTGACCCAGGCGCCCGTCTCGGTGGCAGGAGGCGCGGGAGCG 163

Db 61 CCAGGACTGCCTGACCCAGGCGCCCGTCTCGGTGGCAGGAGGCGCGGGAGCG 120

QY	704	ATCCGAATGCCCAATAATACGTGACCCACAGAGAAGCGCATGATGAGTTT	763
Db	661	ATCCGAATGCCCAATAATACGTGACCCACAGAGAAGCGCATGATGAGTTT	720
QY	764	TTAAGAGGGAACCTGGGAGGACAG	788
Db	721	TTAAGAGGGAACCTGGGAGGACAG	745
RESULT 7			
ABX98011			
ID	ABX98011	standard; cDNA; 1106 BP.	
XX	XX	ABX98011;	
AC	AC		
DT	DT	16-MAY-2003 (first entry)	
XX	XX		
DE	DE	Human PRO polynucleotide #268.	
XX	XX		
KW	KW	Human; PRO; gene; ss; cytosstatic; chromosome mapping; gene mapping;	
KW	KW	protein electrophoresis; tumour necrosis factor-alpha; TNF-alpha; blood;	
XX	XX	chondrocyte differentiation; chondrocyte proliferation; tumour.	
OS	OS	Homo sapiens.	
XX	XX		
PN	PN	US2003032102-A1.	
XX	XX		
PD	PD	13-FEB-2003.	
XX	XX		
PF	PF	17-JUN-2002; 2002US-0173697.	
XX	XX		
PR	PR	16-SEP-1998; 98WO-US19330.	
PR	PR	07-OCT-1998; 98WO-US21141.	
PR	PR	01-DEC-1998; 98WO-US25108.	
PR	PR	08-MAR-1999; 98WO-US05028.	
PR	PR	14-MAY-1999; 98WO-US10733.	
PR	PR	02-JUN-1999; 98WO-US12252.	
PR	PR	01-SEP-1999; 98WO-US20111.	
PR	PR	15-SEP-1999; 98WO-US21090.	
PR	PR	01-DEC-1999; 98WO-US28301.	
PR	PR	02-DEC-1999; 98WO-US28551.	
PR	PR	30-DEC-1999; 98WO-US31274.	
PR	PR	05-JAN-2000; 2000WO-US00219.	
PR	PR	18-FEB-2000; 2000WO-US04341.	
PR	PR	18-FEB-2000; 2000WO-US04342.	
PR	PR	22-FEB-2000; 2000WO-US04414.	
PR	PR	24-FEB-2000; 2000WO-US05004.	
PR	PR	01-MAR-2000; 2000WO-US05601.	
PR	PR	02-MAR-2000; 2000WO-US05841.	
PR	PR	15-MAR-2000; 2000WO-US06884.	
PR	PR	30-MAR-2000; 2000WO-US08439.	
PR	PR	17-MAY-2000; 2000WO-US13705.	
PR	PR	22-MAY-2000; 2000WO-US14042.	
PR	PR	30-MAY-2000; 2000WO-US14941.	
PR	PR	02-JUN-2000; 2000WO-US15264.	
PR	PR	28-JUL-2000; 2000WO-US20710.	
PR	PR	24-AUG-2000; 2000WO-US23328.	
PR	PR	08-NOV-2000; 2000WO-US30952.	
PR	PR	01-DEC-2000; 2000WO-US32678.	
PR	PR	20-DEC-2000; 2000WO-US34956.	
PR	PR	28-FEB-2001; 2001WO-US06520.	
PR	PR	01-JUN-2001; 2001WO-US17800.	
PR	PR	20-JUN-2001; 2001WO-US19692.	
PR	PR	29-JUN-2001; 2001WO-US21066.	
PR	PR	09-JUL-2001; 2001WO-US21735.	
PR	PR	29-AUG-2001; 2001WO-US27099.	
PR	PR	18-SEP-1997; 97US-0592663P.	
PR	PR	18-SEP-1997; 97US-059266P.	
PR	PR	17-OCT-1997; 97US-062250P.	
PR	PR	21-OCT-1997; 97US-063486P.	
PR	PR	24-OCT-1997; 97US-063120P.	
PR	PR	28-OCT-1997; 97US-063512P.	
PR	PR	28-OCT-1997; 97US-063540P.	

PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082688P.
PR 21-APR-1998; 98US-082669P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084619P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085799P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085580P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086332P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087872P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.

PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Query Match 57.4%; Score 743.4; DB 25; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.6e-233;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 44 TGGATCTGGCGGAATGTGGCTGGAGAGTCTGGCGGTACCGTCCAGCTGCC 103
Db 1 TGGATCTGGCGGAATGTGGCTGGAGAGTCTGGCGGTACCGTCCAGCTGCC 60

```
QY 104 CCAGGACTGCCCTGACCCAGCGCCCGCTGCTCGTGGCAGAGAGGCGCGCGAGCG 163
Db 61 CCAGGACTGCCCTGACCCAGCGCGCCCGCTGCTCGTGGCAGAGAGGCGCGCGAGCG 120
QY 164 CCATGGCCTGCATCTGAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGGTTCC 223
Db 121 CCATGGCCTGCATCTGAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGGTTCC 180
QY 224 TTTTCTGCTGGTTGCTGCTGTTGTAATGAAGTGAATTTCCCAATTCCTACTAACTGCT 283
Db 181 TTTTCTGCTGGTTGCTGCTGTTGTAATGAAGTGAATTTCCCAATTCCTACTAACTGCT 240
QY 284 TTGACAACTGGTCAAAAGTGAATACCATCTTCCTAACATCAGCGCGCCCTTCGAA 343
Db 241 TTGACAACTGGTCAAAAGTGAATACCATCTTCCTAACATCAGCGCGCCCTTCGAA 300
QY 344 CTCACATGATGATACATAATGTAAGACACAAGAGCCTTTGCAACTGGACTGTGACCTTT 403
Db 301 CTCACATGATGATACATAATGTAAGACACAAGAGCCTTTGCAACTGGACTGTGACCTTT 360
QY 404 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAAATGAGATAGTC 463
Db 361 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAAATGAGATAGTC 420
QY 464 GNTCCTCCTGCATTTGGAGATGACAAATGCTCCGCCACCAAGGTTATGAGAGATGTCG 523
Db 421 GNTCCTCCTGCATTTGGAGATGACAAATGCTCCGCCACCAAGGTTATGAGAGATGTCG 480
QY 524 GCSCATGACCATGATTCAGTGTGTCCTCATACAGAGGTTCCCTTTTCTGCTAAAAACC 583
Db 481 GCSCATGACCATGATTCAGTGTGTCCTCATACAGAGGTTCCCTTTTCTGCTAAAAACC 540
QY 584 CTGATTTATTTTCAGGAAGCGAATACATCTATTTGTTGTTGTTGGGACCTTTCGCA 643
Db 541 CTGATTTATTTTCAGGAAGCGAATACATCTATTTGTTGTTGTTGGGACCTTTCGCA 600
QY 644 ATATGAGGAAGATGGCAATGGCATCGTTTACAACTGTTGAAAAGACAGTGTGATCT 703
Db 601 ATATGAGGAAGATGGCAATGGCATCGTTTACAACTGTTGAAAAGACAGTGTGATCT 660
QY 704 ATCCGAATGCCAAATATACGTGACACACAGAGCGCATGAGTTACTGTGATGGAGTTT 763
Db 661 ATCCGAATGCCAAATATACGTGACACACAGAGCGCATGAGTTACTGTGATGGAGTTT 720
QY 764 TTAGAGAGGAAGTGGGAGGACAG 788
Db 721 TTAGAGAGGAAGTGGGAGGACAG 745

RESULT 8
ABX78795
ID ABX78795 standard; cDNA; 1106 BP.
XX AC ABX78795;
XX XX
DT 15-APR-2003 (first entry)
XX XX
DE Human PRO polynucleotide #268.
XX XX
KW Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach;
KW liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX OS Homo sapiens.
XX XX
PN US2003027272-A1.
XX XX
PD 06-FEB-2003.
XX XX
PF 21-JUN-2002; 2002US-0176492.
XX XX
PR 15-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-06250P.
PR 21-OCT-1997; 97US-06348P.
PR 24-OCT-1997; 97US-063120P.
PR 28-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 11-DEC-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069333P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081836P.
```

PR	26-JUN-1998;	98US-09101010P.
PR	01-JUL-1998;	98US-091359P.
PR	01-JUL-1998;	98US-091544P.
PR	02-JUL-1998;	98US-091478P.
PR	02-JUL-1998;	98US-091486P.
PR	02-JUL-1998;	98US-091626P.
PR	02-JUL-1998;	98US-091628P.
PR	02-JUL-1998;	98US-091632P.
PR	24-JUL-1998;	98US-094006P.
PR	04-AUG-1998;	98US-095282Z.
PR	10-AUG-1998;	98US-095998P.
PR	10-AUG-1998;	98US-096012P.
PR	17-AUG-1998;	98US-096757P.
PR	17-AUG-1998;	98US-096766P.
PR	17-AUG-1998;	98US-096867P.
PR	17-AUG-1998;	98US-096891P.
PR	17-AUG-1998;	98US-096897P.
PR	18-AUG-1998;	98US-096949P.
PR	18-AUG-1998;	98US-096959P.
PR	18-AUG-1998;	98US-097022P.
PR	26-AUG-1998;	98US-097952P.
PR	26-AUG-1998;	98US-097954P.
PR	26-AUG-1998;	98US-097955P.
PR	26-AUG-1998;	98US-097971P.
PR	26-AUG-1998;	98US-097974P.
PR	01-SEP-1998;	98US-098014P.
PR	01-SEP-1998;	98US-098176P.
PR	01-SEP-1998;	98US-098723P.
PR	02-SEP-1998;	98US-098803P.
PR	02-SEP-1998;	98US-098821P.
PR	02-SEP-1998;	98US-098843P.
PR	09-SEP-1998;	98US-099602P.
PR	10-SEP-1998;	98US-099741P.
PR	10-SEP-1998;	98US-099746P.

Qy	44	TGATCTCGGGAATGTGGCTGGAGAGGTCCTGGCCGTGGTGTTACACGCTTCCAGAGCTGCC	103
Db	1	TGATCTCGGGAATGTGGCTGGAGAGGTCCTGGCCGTGGTGTTACACGCTTCCAGAGCTGCC	60
Qy	104	CCAGAGACTGCCCTTGACCCACGCGCGCCGCTCGTCTCGTGGCAGAGAGGCGCGGCGAGCG	163
Db	61	CCAGGACTGCCCTTGACCCACGCGCGCCGCTCGTCTCGTGGCAGAGAGGCGCGGCGAGCG	120
Qy	164	CCATGGCCTGCATCCTGCAAGAGAAAGTCGTGTGATTCCTGTGAGCTTCATAGCAGGCTTCC	223
Db	121	CCATGGCCTGCATCCTGAAGAGAAAGTCGTGTGATTCCTGTGAGCTTCATAGCAGGCTTCC	180
Qy	224	TTTTCTCTGCTGGTTGTGGCTCTGTGTAATGAAGTGAATTTGCCATTTGCCATTTAAACTGCT	283
Db	181	TTTTCTCTGCTGGTTGTGGCTCTGTGTAATGAAGTGAATTTGCCATTTGCCATTTAAACTGCT	240
Qy	284	TTGGACACACTGGTACAAAGTGGATPACCATTTCTTCCTACACATACAGGCGGCCCCCTTGGAA	343
Db	241	TTGGACACACTGGTACAAAGTGGATPACCATTTCTTCCTACACATACAGGCGGCCCCCTTGGAA	300
Qy	344	CTCAGTATGGATACATAAATGTGAAGACACAAAGAGCCTTTTGCAACTGGACTGTGACCTTT	403
Db	301	CTCAGTATGGATACATAAATGTGAAGACACAAAGAGCCTTTTGCAACTGGACTGTGACCTTT	360
Qy	404	GTGCCATATGTCAAACTCAGCTCAGATGGTTGGCCAGAGAGTGGGAATAGAGATGATC	463
Db	361	GTGCCATATGTCAAACTCAGCTCAGATGGTTGGCCAGAGAGTGGGAATAGAGATGATC	420
Qy	464	GATCTCTCTGCAATTTGGAGAAATGAACAATGCCCCACCAAAAGTTATGAAAGAAAGATGTCG	523
Db	421	GATCTCTCTGCAATTTGGAGAAATGAACAATGCCCCACCAAAAGTTATGAAAGAAAGATGTCG	480
Qy	524	GCCGCATGACCATGATTCGAGTTGTGTCCCATACCAAGGCTTCTCTTTTGTCTAAAAAACC	583
Db	481	GCCGCATGACCATGATTCGAGTTGTGTCCCATACCAAGGCTTCTCTTTTGTCTAAAAAACC	540

QY	584	CTGATTATTTTCAAGGAAGCACTACTATTCTGTTATTGTTGGGACCTTTCCGCA	643	PR	20-DEC-1999;	99WO-US30999.
Dd	541		600	PR	22-DEC-1999;	99WO-US30720.
QY	644	ATATGAGGAAGATGCGATGCGTTTACACATGTTGAAAGACAGTTGTATCT	703	PR	30-DEC-1999;	99WO-US31243.
Dd	601	ATATGAGGAAGATGCGATGCGTTTACACATGTTGAAAGACAGTTGTATCT	660	PR	30-DEC-1999;	99WO-US31274.
QY	704	ATCCGAATGCCCAATATAGTGAACACAGAGGACGCGATGTTACTGTGATGAGTTT	763	PR	05-JAN-2000;	2000WO-US00219.
Dd	661		720	PR	06-JAN-2000;	2000WO-US00277.
QY	764	TTAAGAAGAACTGGGAAGACAG	788	PR	11-FEB-2000;	2000WO-US03376.
Dd	721		745	PR	18-FEB-2000;	2000WO-US03565.
RESULT 9				PR	18-FEB-2000;	2000WO-US04341.
ABX75808				PR	18-FEB-2000;	2000WO-US04342.
ID	ABX75808	standard; cDNA; l106 BP.		PR	22-FEB-2000;	2000WO-US04414.
AC	ABX75808;			PR	24-FEB-2000;	2000WO-US05004.
XX				PR	01-MAR-2000;	2000WO-US05601.
DT	26-MAR-2003	(first entry)		PR	02-MAR-2000;	2000WO-US05841.
DE		Human cDNA encoding secreted/transmembrane protein, PRO17177.		PR	10-MAR-2000;	2000WO-US06319.
KW		Human; ss: gene; secreted protein; transmembrane protein; PRO;		PR	15-MAR-2000;	2000WO-US06884.
KW		antiarthritic; vulnery; tumour necrosis factor-alpha;		PR	21-MAR-2000;	2000WO-US07532.
KW		chondrocyte cell proliferation; chondrocyte cell differentiation;		PR	30-MAR-2000;	2000WO-US08439.
KW		tumour; adrenal tumour; lung tumour; colon tumour; breast tumour;		PR	17-MAY-2000;	2000WO-US13705.
KW		prostate tumour; rectal tumour; cervical tumour; liver tumour;		PR	22-MAY-2000;	2000WO-US14042.
KW		bone disorder; cartilage disorder; arthritis; sports injury.		PR	30-MAY-2000;	2000WO-US14941.
OS		Homo sapiens.		PR	02-JUN-2000;	2000WO-US15284.
XX				PR	28-JUL-2000;	2000WO-US20710.
PN	US2003022298-A1.			PR	11-AUG-2000;	2000WO-US22031.
XX				PR	23-AUG-2000;	2000WO-US23522.
PD	30-JAN-2003.			PR	24-AUG-2000;	2000WO-US23328.
PF				PR	08-NOV-2000;	2000WO-US30952.
XX				PR	10-NOV-2000;	2000WO-US30873.
XX				PR	01-DEC-2000;	2000WO-US32678.
XX				PR	20-DEC-2000;	2000WO-US34956.
XX				PR	28-FEB-2001;	2001WO-US06520.
XX				PR	22-MAR-2001;	2001WO-US09552.
XX				PR	25-MAY-2001;	2001WO-US17092.
XX				PR	01-JUN-2001;	2001WO-US17800.
XX				PR	20-JUN-2001;	2001WO-US19692.
XX				PR	29-JUN-2001;	2001WO-US21066.
XX				PR	09-JUL-2001;	2001WO-US21735.
XX				PR	29-AUG-2001;	2001WO-US27099.
XX				PR	18-SEP-1997;	97US-059263P.
XX				PR	18-SEP-1997;	97US-059266P.
XX				PR	17-OCT-1997;	97US-062250P.
XX				PR	21-OCT-1997;	97US-063486P.
XX				PR	24-OCT-1997;	97US-063120P.
XX				PR	28-OCT-1997;	97US-063540P.
XX				PR	28-OCT-1997;	97US-063544P.
XX				PR	28-OCT-1997;	97US-063544P.
XX				PR	28-OCT-1997;	97US-063564P.
XX				PR	29-OCT-1997;	97US-063734P.
XX				PR	31-OCT-1997;	97US-063870P.
XX				PR	31-OCT-1997;	97US-064103P.
XX				PR	13-NOV-1997;	97US-065311P.
XX				PR	21-NOV-1997;	97US-066120P.
XX				PR	24-NOV-1997;	97US-066466P.
XX				PR	24-NOV-1997;	97US-066772P.
XX				PR	11-DEC-1997;	97US-069335P.
XX				PR	12-DEC-1997;	97US-069425P.
XX				PR	17-DEC-1997;	97US-069870P.
XX				PR	18-DEC-1997;	97US-068017P.
XX				PR	10-MAR-1998;	98US-077450P.
XX				PR	11-MAR-1998;	98US-077632P.
XX				PR	11-MAR-1998;	98US-077649P.
XX				PR	20-MAR-1998;	98US-078866P.
XX				PR	20-MAR-1998;	98US-078939P.
XX				PR	27-MAR-1998;	98US-079664P.
XX				PR	27-MAR-1998;	98US-079786P.
XX				PR	31-MAR-1998;	98US-080107P.
XX				PR	31-MAR-1998;	98US-080194P.
XX				PR	01-APR-1998;	98US-080327P.
XX				PR	01-APR-1998;	98US-080333P.
XX				PR	08-APR-1998;	98US-081049P.
XX				PR	08-APR-1998;	98US-081070P.
XX				PR	09-APR-1998;	98US-081195P.

PR	26-JUN-1998;	98US-090863P.
Query Match	57.4%;	Score 743.4; DB 25; Length 1106;
Best Local Similarity	99.9%;	Pred. No. 4.6e-233;
Matches 744;	Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	44	TGATCTGGGGAACTGGGCTGGAGAGTCTTCGTGGTTACAGCCTCACGCCTGCC 103
Db	1	TGATCTGGGGAACTGGGCTGGAGAGTCTTCGTGGTTACAGCCTCACGCCTGCC 60
QY	104	CCAGACTGCCCTTGACCACGAGCGCGCGTCTCGCTGGCAGAGAGCGCGGGAGCG 163
Db	61	CCAGACTGCCCTTGACCACGAGCGCGCGTCTCGCTGGCAGAGAGCGCGGGAGCG 120
QY	164	CCATGGCCTGCATPCCTGAAGAAGTCTGTGATTGCTGTGAGCTTCATAGACGCTTC 223
Db	121	CCATGGCCTGCATPCCTGAAGAAGTCTGTGATTGCTGTGAGCTTCATAGACGCTTC 180
QY	224	TTTTCTCTGCTGGTGGCGTCTGTAAATGAAGTGAATTTCCCATGCTACTAACTGCT 283
Db	181	TTTTCTCTGCTGGTGGCGTCTGTAAATGAAGTGAATTTCCCATGCTACTAACTGCT 240
QY	284	TTGGACAACCTGGTACAAAAGTGATACCACTTCCTACACATACAGGCGGCCCTTGAA 343
Db	241	TTGGACAACCTGGTACAAAAGTGATACCACTTCCTACACATACAGGCGGCCCTTGAA 300
QY	344	CTCACTATGATACATAAATGTGAAGACACAGAGCCTTTGCAACTGACCTGTGACTGT 403
Db	301	CTCACTATGATACATAAATGTGAAGACACAGAGCCTTTGCAACTGACCTGTGACTGT 360
QY	404	GTCGCATAGTGTCAAACCTCAGCTCAGATGTTGGCCAGAGGTGGGAAATGAGATAGTC 463
Db	361	GTCGCATAGTGTCAAACCTCAGCTCAGATGTTGGCCAGAGGTGGGAAATGAGATAGTC 420
QY	464	GATCCTCTGCATTTGGACAATGACAAATGCCGCCACCAAGGTTATGAAGAAGATGTCG 523
Db	421	GATCCTCTGCATTTGGAGAATGAACAATGCCGCCACCAAGGTTATGAAGAAGATGTCG 480
QY	524	GCGCGATGACCATGATTCGAGTTGTGCCATACAGCGTTCTCTTTGCTAAAAAAC 583
Db	481	GCGCGATGACCATGATTCGAGTTGTGCCATACAGCGTTCTCTTTGCTAAAAAAC 540
QY	584	CTGATTAATTTTCAAGAGAGCAATACTATTTGTATTATTTGGGAGACTTTCGCA 643
Db	541	CTGATTAATTTTCAAGAGAGCAATACTATTTGTATTATTTGGGAGACTTTCGCA 600
QY	644	ATATGAGGAAGATGCAATGGCATGCTGTTWACACATGTTGAAAAGACAGTTGGTATCT 703
Db	601	ATATGAGGAAGATGCAATGGCATGCTGTTWACACATGTTGAAAAGACAGTTGGTATCT 660
QY	704	ATCCGAATGCCAAATATATACGTGACCACAGAGAGCGCATGASTTACTGTGATGGAGTTT 763
Db	661	ATCCGAATGCCAAATATATACGTGACCACAGAGAGCGCATGASTTACTGTGATGGAGTTT 720
QY	764	TTAAGAGGAAGTGGGAGGACAG 788
Db	721	TTAAGAGGAAGTGGGAGGACAG 745
RESULT 10		
ABX77013		
ID	ABX77013 standard;	CDNA; 1106 BP.
XX		
AC	ABX77013;	
XX		
DT	04-APR-2003 (first entry)	
XX		
DE	Human PRO polynucleotide #268.	
XX		
KW	Human; PRO; gene; ss; tumour necrosis factor-alpha; blood; cancer;	
KW	chondrocyte cell; tumour; adrenal tumour; lung; colon; breast; prostate;	
KW	kidney; rectum; cervix; liver; bone disorder; cartilage disorder;	
KW	arthritis; sports injury; genetic disorder; antithrptic; vulnervary.	

XX OS Homo sapiens.
XX PN US2003027280-AL.
XX PD 06-FEB-2003.
XX PF 20-JUN-2002; 2002US-0176993.
XX PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US04414.
PR 22-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13709.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084840P.
PR 07-MAY-1998; 98US-084843P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 05-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.

Db 541 CCGATATATTTTCAAGGAAGGAATACTACTATTATGTTATTTGGGAACCTTCCGCA 600
Qy 644 ATATGAGGAAGATGGCAATGGATCGTTTACACATGTTGAAAGAGACAGTTGTATCT 703
Db 601 ATATGAGGAAGATGGCAATGGATCGTTTACACATGTTGAAAGAGACAGTTGTATCT 660
Qy 704 ATCCGAATGCCCAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 763
Db 661 ATCCGAATGCCCAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 720
Qy 764 TTAAGAGGAAGAACTGGGAAGACAG 788
Db 721 TTAAGAGGAAGAACTGGGAAGACAG 745
RESULT 12
AAD09952
ID AAD09952 standard; cDNA; 1122 BP.
XX AAD09952;
XX
PT 12-SEP-2001 (first entry)
XX
DE Human drug metabolising enzyme (DME-17) cDNA.
XX
KW Human; drug metabolising enzyme; DME-17; immunosuppressive; gene therapy;
KW cytosolic; autoimmune disorder; inflammatory disorder; atherosclerosis;
KW osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism;
KW rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia;
KW developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy;
KW thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma;
KW gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma;
KW actinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma;
KW cell proliferative disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 123..755
FT /*tag= a
FT /product= "Human drug metabolising enzyme (DME-17)"
FT sig_peptide 123..212
FT /*tag= b
FT mat_peptide 213..752
FT /*tag= c
FT /product= "Mature drug metabolising enzyme (DME-17)"
XX
PN WO200151638-A2.
XX
XX 19-JUL-2001.
XX
XX 12-JAN-2001; 2001WO-US01174.
XX
PR 14-JAN-2000; 2000US-0176139.
PR 21-JAN-2000; 2000US-0177443.
PR 28-JAN-2000; 2000US-0178574.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI Ring HZ, Hillman JL, Yue H, Azimzai Y, Yao MG, Gandhi AR;
PI Nguyen DB, Tang YT, Lal P, Bandwan O;
XX
DR WPI: 2001-425874/45.
DR P-PSDB; AAE05186.
XX
PT Drug metabolizing enzymes and encoding polynucleotides, useful for
PT diagnosing, treating and/or preventing autoimmune, inflammatory, cell
PT proliferative, developmental, endocrine, eye, metabolic, and
PT gastrointestinal disorders -
XX
PT Claim 5; Page 168; 133pp; English.

XX The present sequence is human drug metabolising enzyme (DME-17) cDNA.
CC Human DME and its nucleic acid molecule are useful for the diagnosis,
CC treatment and prevention of disorders associated with increased or
CC decreased expression of DME. Examples of such disorders include,
CC autoimmune/inflammatory disorder such as acquired immune deficiency
CC syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative
CC disorder such as actinic keratosis, atherosclerosis; developmental
CC disorder such as epilepsy, anaemia; endocrine disorder such as
CC acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as
CC diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis;
CC metabolic disorder such as Addison's disease, obesity; gastrointestinal
CC disorder such as anorexia, dysphagia and hepatic tumours including
CC nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for
CC creating 'knockin' humanised animals (pigs) or transgenic animals (mice
CC or rats) to model human disease. DME DNA is also in useful is gene
CC therapy. DME and its immunogenic fragments are useful for screening
CC libraries of compounds in several drug screening assays.
XX
SQ Sequence 1122 BP; 338 A; 228 C; 261 G; 295 T; 0 other;
Query Match 57.4%; Score 743.4; DB 22; Length 1122;
Best Local Similarity 99.9%; Pred. No. 4.7e-233;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 44 TGGATCTGCGGGAATGCTGGCTGGAGAGTCTCTGCCGTGTACCGACCTCCAGCTGCC 103
Db 1 TGGATCTGCGGGAATGCTGGCTGGAGAGTCTCTGCCGTGTACCGACCTCCAGCTGCC 60
Qy 104 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGCGGAGCG 163
Db 61 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGCGGAGCG 120
Qy 164 CCATGSCCTGATCTCTGAGAGAGAAAGTCTGTGATCTGTGAGCTTCTATACAGCTTC 223
Db 121 CCATGSCCTGATCTCTGAGAGAGAAAGTCTGTGATCTGTGAGCTTCTATACAGCTTC 180
Qy 224 TTTTCTGCTGTTGCTGCTTGTAAATGAAGTAATTTCCCATGCTTACTAACTGCT 283
Db 181 TTTTCTGCTGTTGCTGCTTGTAAATGAAGTAATTTCCCATGCTTACTAACTGCT 240
Qy 284 TTGGACAACCTGGTACAAAAGTGGATACCATTTCTCTACACATACAGCGGCCCTTCGAA 343
Db 241 TTGGACAACCTGGTACAAAAGTGGATACCATTTCTCTACACATACAGCGGCCCTTCGAA 300
Qy 344 CTCCTAATGATACATAAATGTGAAGACACAGAGCCTTTCGAACTGCTGACCTTT 403
Db 301 CTCCTAATGATACATAAATGTGAAGACACAGAGCCTTTCGAACTGCTGACCTTT 360
Qy 404 GTGCCATAGTGTCAAACTCAGGTGATGTTGCCAGAGGTGGGAAATGAGATAGATC 463
Db 361 GTGCCATAGTGTCAAACTCAGGTGATGTTGCCAGAGGTGGGAAATGAGATAGATC 420
Qy 464 GATCCTCTGCTGTTGGAGAAATGAACAATGCCGCCCAAGGTTATGAAGAATGTG 523
Db 421 GATCCTCTGCTGTTGGAGAAATGAACAATGCCGCCCAAGGTTATGAAGAATGTG 480
Qy 524 GCCCATGACCATGATTCGAGTTGTGCTCCCATACAGCTTCTCTTTTGTAAAAAACC 583
Db 481 GCCCATGACCATGATTCGAGTTGTGCTCCCATACAGCTTCTCTTTTGTAAAAAACC 540
Qy 584 CTGATTAATTTTCAAGGAAGCAATACTACTATTTGTTTATTTGGGACCTTTCGCA 643
Db 541 CTGATTAATTTTCAAGGAAGCAATACTACTATTTGTTTATTTGGGACCTTTCGCA 600
Qy 644 ATATGAGGAAGATGGCAATGGATCGTTTACACATGTTGAAAGAGACAGTTGTATCT 703
Db 601 ATATGAGGAAGATGGCAATGGATCGTTTACACATGTTGAAAGAGACAGTTGTATCT 660
Qy 704 ATCCGAATGCCCAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 763
Db 661 ATCCGAATGCCCAATATACGTGACACAGAGAGCGCATGAGTTACTGTGATGAGTTT 720

QY 764 TTAGAGGAACTGGGAGGACAG 788
 DB 721 TTAAGAAGAACTGGGAGGACAG 745

RESULT 13
 AAD34123
 ID AAD34123 standard; cDNA; 2575 BP.
 XX AAD34123;
 AC AAD34123;
 XX 16-JUL-2002 (first entry)
 XX Human secreted protein-encoding gene 1 cDNA clone HE9PR39, SEQ ID NO:28.
 XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
 KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
 KW diabetes; infection; wound healing; vulnerability; chemotaxis; food additive;
 KW gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 211..1128
 FT /tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 211..300
 FT /tag= b
 FT mat_peptide 301..1125
 FT /tag= c
 FT /product= "Mature secreted protein"
 XX WO200222654-A1.
 PN
 XX 21-MAR-2002.
 PD
 XX 17-JAN-2001; 2001WO-US01385.
 PF
 XX 12-SEP-2000; 2000US-231969P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Sopet DR, Olsen HS;
 PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
 PI Ni J;
 XX WPI; 2002-315684/35.
 DR P-PSDB; AAE21595.
 XX Seventeen nucleic acid molecules encoding human secreted proteins,
 PT useful for treating and preventing cancer, immune disorders (e.g.
 PT Addison's disease, and allergies), and cardiovascular disorders (e.g.
 PT myocardial ischemias) -
 XX Claim 1; Page 423-424; 483pp; English.
 PS
 XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted
 CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
 CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 17 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of cancer,
 CC proliferative disorders, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (rheumatoid arthritis), inflammation, allergies, neurological

CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
 CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
 CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney
 CC disorders, gastrointestinal disorders, pregnancy-related disorders,
 CC endocrine disorders, and infections. The proteins can also be used to aid
 CC wound healing and epithelial cell proliferation, to prevent skin aging
 CC due to sunburn, to maintain organs before transplantation, for supporting
 CC cell culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties. The
 CC present sequence represents a human secreted protein-encoding cDNA of
 CC the invention.
 XX
 SQ Sequence 2575 BP; 758 A; 541 C; 521 G; 752 T; 3 other;

Query Match 55.2%; Score 714; DB 24; Length 2575;
 Best Local Similarity 98.6%; Pred. No. 3.8e-223;
 Matches 720; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 65 TGGAGAGTCTCCCGTGGTACACAGCTCCAGCCGCCGCCAGAGCTGCCCTGACCCAG 124
 DB 110 TCGGCAGGAGCTCCCGTGGTACACAGCTCCAGCCGCCGCCAGAGCTGCCCTGACCCAG 169
 QY 125 GCGGCCCGCTCTCGGTGCGAGAGCGCGCGGAGCGCCATGGCCTGCATCCTGAAGA 184
 DB 170 GCGGCCCGCTCTCGGTGCGAGAGCGCGCGGAGCGCCATGGCCTGCATCCTGAAGA 229
 QY 185 GAAAGTCTGTGATTCCTGTGAGCTTCATGACAGCGTTCCTTTCCTGCTGGTGGCGTC 244
 DB 230 GAAAGTCTGTGATTCCTGTGAGCTTCATGACAGCGTTCCTTTCCTGCTGGTGGCGTC 289
 QY 245 TTGTAATGAAGTGAATTCCTGCTACTAACTGCTTTGGACAACCTGGTACAAGT 304
 DB 290 TTGTAATGAAGTGAATTCCTGCTACTAACTGCTTTGGACAACCTGGTACAAGT 349
 QY 305 GGATACCATTCCTPACACATACAGCGCGCCCTTCGAACTCACTATGATPACATAAATG 364
 DB 350 GGATACCATTCCTPACACATACAGCGCGCCCTTCGAACTCACTATGATPACATAAATG 409
 QY 365 TGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTTGTGCCATAGTCTCAAACTCAG 424
 DB 410 TGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTTGTGCCATAGTCTCAAACTCAG 469
 QY 425 GTCAGATGTTGGCCAGAGGTGGAAATGAGATAGATGATCCTCCTGCTATTGGAGAA 484
 DB 470 GTCAGATGTTGGCCAGAGGTGGAAATGAGATAGATGATCCTCCTGCTATTGGAGAA 529
 QY 485 TGAACATGCCCCCACCACCAAGGTTATGAAGAAGATGTGGCCGCATGACCATGATTCAG 544
 DB 530 TGAACATGCCCCCACCACCAAGGTTATGAAGAAGATGTGGCCGCATGACCATGATTCAG 589
 QY 545 TTGTGTCCCATACACAGCTTCCTCTTTGCTAAAAAACCTGATTATTTTTCAGGAAG 604
 DB 590 TTGTGTCCCATACACAGCTTCCTCTTTGCTAAAAAACCTGATTATTTTTCAGGAAG 649
 QY 605 CGAATACTACTATTTGTTATTTTGGGACCTTTCCGCAATATGAGGAAGATGCGCAATG 664
 DB 650 CGAATACTACTATTTGTTATTTTGGGACCTTTCCGCAATATGAGGAAGATGCGCAATG 709
 QY 665 GCATCTTTTACACATGTTGAAAGACAGCTGGTATCTATCCGAATGCCCAATATAG 724
 DB 710 GCATCTTTTACACATGTTGAAAGACAGCTGGTATCTATCCGAATGCCCAATATAG 769
 QY 725 TGACCCACAGAGAGCGCATGAGTTACTGTGAGGAGTTTTTAAGAGAGAACTGGGAAG 784
 DB 770 TGACCCACAGAGAGCGCATGAGTTACTGTGAGGAGTTTTTAAGAGAGAACTGGGAAG 829
 QY 785 ACAGGGGGCA 794
 DB 830 ACAGAGTCCA 839

RESULT 14

AAD34106
 ID AAD34106 standard; cDNA; 3189 BP.
 XX
 AC AAD34106;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Human secreted protein-encoding gene 1 cDNA clone HE9PR39, SEQ ID NO:1.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KW immune system; cardiovascular; angioenic; gastrointestinal; endocrine;
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
 KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
 KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
 KW gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 100..1017
 FT /tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 100..189
 FT /tag= b
 FT mat_peptide 190..1014
 FT /tag= c
 FT /product= "Mature secreted protein"
 XX
 PN WO200222654-A1.
 XX
 XX 21-MAR-2002.
 PD
 XX 17-JAN-2001; 2001WO-US01385.
 PF
 XX 12-SEP-2000; 2000US-231969P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
 PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
 PI Ni J;
 XX
 DR WPI: 2002-315684/35.
 DR P-PSDB; AAE21578.
 XX
 PT Seventeen nucleic acid molecules encoding human secreted proteins,
 PT useful for treating and preventing cancer, immune disorders (e.g.,
 PT Addison's disease, and allergies), and cardiovascular disorders (e.g.,
 PT myocardial ischemias)
 XX
 PS Claim 1; Page 413-414; 483pp; English.
 XX
 CC AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted
 CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
 CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 17 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of cancer,
 CC proliferative disorders, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
 CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
 CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
 CC atherosclerosis, cardiovascular disorders, angioenic disorders, kidney
 CC disorders, gastrointestinal disorders, pregnancy-related disorders,
 CC endocrine disorders, and infections. The proteins can also be used to aid
 CC wound healing and epithelial cell proliferation, to prevent skin aging

CC due to sunburn, to maintain organs before transplantation, for supporting
 CC cell culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties. The
 CC present sequence represents a human secreted protein-encoding cDNA of
 CC the invention.
 XX
 SQ Sequence 3189 BP; 993 A; 614 C; 624 G; 958 T; 0 other;
 Query Match 55.28; Score 713.8; DB 24; Length 3189;
 Best Local Similarity 99.08; Pred. No. 5e-223;
 Matches 718; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 70 AGTGCTCTCGTGGTGTGACCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 249
 DB 4 ACGAGCTGCGGTGTGACCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 183
 QY 130 CCGCTGCTCTCGTGGTGTGACCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 189
 DB 64 CCGCTGCTCTCGTGGTGTGACCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 123
 QY 190 TCTGTGATTGCTGTGAGCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 249
 DB 124 TCTGTGATTGCTGTGAGCTTCATAGCAGCGTTCCCTTCTCTGCTGGTGTGCGTCTTCTA 183
 QY 250 AATGAAGTGAATTTCCCAATTCCTTAACTGCTTGGACAACCTGGTACAAAGTGGATA 309
 DB 184 AATGAAGTGAATTTCCCAATTCCTTAACTGCTTGGACAACCTGGTACAAAGTGGATA 243
 QY 310 CCATTCTCTACATACATAGCGCGCCCTTCGAACTCACTATGATACATATAATGTGAAG 369
 DB 244 CCAATCTCTACATACATAGCGCGCCCTTCGAACTCACTATGATACATATAATGTGAAG 303
 QY 370 ACACAAGAGCCCTTTCACACTGCTGACCTTTGCGCATAGTGTCAAACCTCAGGTGAG 429
 DB 304 ACACAAGAGCCCTTTCACACTGCTGACCTTTGCGCATAGTGTCAAACCTCAGGTGAG 363
 QY 430 ATGTTGGCCAGAGGTTGGAATGAGATAGATCGATCCCTCTGCTGATTTGGAGAATGAAC 489
 DB 364 ATGTTGGCCAGAGGTTGGAATGAGATAGATCGATCCCTCTGCTGATTTGGAGAATGAAC 423
 QY 490 AATGCCCCCAACAAAGTTATGAAGAAGATGTGCGCCGCGCATGACCATGATTCGAGTTGTG 549
 DB 424 AATGCCCCCAACAAAGTTATGAAGAAGATGTGCGCCGCGCATGACCATGATTCGAGTTGTG 483
 QY 550 TCCCATACAGCGTTCTCTTTTCTTAAACCTGATTTATTTTTCAGGAAGCGGAAT 609
 DB 484 TCCCATACAGCGTTCTCTTTTCTTAAACCTGATTTATTTTTCAGGAAGCGGAAT 543
 QY 610 ACTACTATTGTTGTTATTTGGGACCTTTCCGCAATATGAGGAAGATGGAATGGCATC 669
 DB 544 ACTACTATTGTTATTTGGGACCTTTCCGCAATATGAGGAAGATGGAATGGCATC 603
 QY 670 GTTTACACATCTTGAAGAAAGACAGTTGGTATCTATCCGAATGCCAAATATACGTGACC 729
 DB 604 GTTTACACATCTTGAAGAAAGACAGTTGGTATCTATCCGAATGCCAAATATACGTGACC 663
 QY 730 ACGAGAACCCATCAGTACTGCTGATGAGATTTTAAAGAGAACTGGGAGGACGAGG 789
 DB 664 ACGAGAACCCATCAGTACTGCTGATGAGATTTTAAAGAGAACTGGGAGGACGAGG 723
 QY 790 GGGCA 794
 DB 724 GTCCA 728
 RESULT 15
 AAE21578
 ID AAE21578 standard; cDNA; 696 BP.
 XX
 AC AAE21578;
 XX
 DT 08-FEB-2001 (first entry)

xx Human ORFX ORF2033 polynucleotide sequence SEQ ID NO:4065.
DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
xx vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
xx anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
xx immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
xx hypotensive; dermatological; immunosuppressive; antiinflammatory;
xx antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
xx antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
xx neurodegenerative disorder; osteoarthritis; graft vs host disease;
xx cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
xx cholesterol ester storage; systemic lupus erythematosus; infection;
xx severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
xx allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
xx bone damage; cartilage damage; antiinflammatory disease; coagulation;
xx thrombosis; contraceptive; ss.
OS Homo sapiens.
xx W0200058473-A2.
xx 05-OCT-2000.
xx 31-MAR-2000; 2000WO-US08621.
xx 31-MAR-1999; 99US-0127607.
xx 02-APR-1999; 99US-0127636.
xx 03-APR-1999; 99US-0127728.
xx 30-MAR-2000; 2000US-0540763.
xx (CURA-) CURAGEN CORP.
xx Shimkets RA, Leach M;
xx WPI; 2000-602362/57.
xx P-PSDB; AAB42269.
xx Novel nucleic acids and peptides derived from open reading frame X,
xx useful for treating e.g. cancers, proliferative disorders,
xx neurodegenerative disorders and cardiovascular disease -
xx Claim 5; Page 3256; 5507pp; English.
xx AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
xx which represent the human ORFX open reading frames 1 to 3161. The ORFX
xx sequences have activities such as: cytostatic; hepatotropic; vulnary;
xx antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
xx osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
xx immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
xx antidiabetic; hypotensive; dermatological; immunosuppressive;
xx antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
xx antithyroid; and antianaemic. The sequences can be used for determining
xx the presence of or predisposition to, or preventing or treating
xx pathological conditions associated with an ORFX-associated disorder. The
xx nucleic acids can be used to express ORFX proteins in gene therapy
xx vectors. The proteins and nucleic acids may be used to treat cancers,
xx proliferative disorders, neurodegenerative disorders, osteoarthritis,
xx graft vs host disease, cardiovascular disease, diabetes mellitus,
xx hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
xx erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
xx bacterial or fungal infection, malaria, autoimmune disorders, asthma,
xx allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
xx nocturnal haemoglobinuria, antiinflammatory disease; to enhance
xx coagulation; to inhibit thrombosis; and as a contraceptive.
xx Sequence 696 BP; 191 A; 148 C; 175 G; 179 T; 3 other;

Query Match 50.48; Score 652.4; DB 21; Length 696;
Best Local Similarity 98.8%; Pred. No. 2.8e-203;
Matches 656; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 125 GCGCGCCGCTGCTGCTGAGAGGGCGCGGAGCGCCATGCGTCATCCTGAAGA 184

Db 2 GCGCGCCGCTGCTGCTGAGAGGGCGCGGAGCGCCATGCGTCATCCTGAAGA 61
QY 185 GAAAGTCTGTGATTCCTGTGAGCTTCATAGACGGTTCCTTTCTGCTGGTGGCGTC 244
Db 62 GAAAGTCTGTGATTCCTGTGAGCTTCATAGACGGTTCCTTTCTGCTGGTGGCGTC 121
QY 245 TTGTAAATGAGTGAATTCCTGCTAACTGCTTTGGACACCTGGTACAAAGT 304
Db 122 TTGTAAATGAGTGAATTCCTGCTAACTGCTTTGGACACCTGGTACAAAGT 181
QY 305 GGATACCATTCCTACACATACAGCGCGCCCTTCGAACCTCACTATGATACATAAATG 364
Db 182 GGATACCATTCCTACACATACAGCGCGCCCTTCGAACCTCACTATGATACATAAATG 241
QY 365 TGAAGACACAGAGCCCTTGGCACTGGACTGTGACCTTTGTGCCATAGTGTCAAACTCAG 424
Db 242 TGAAGACACAGAGCCCTTGGCACTGGACTGTGACCTTTGTGCCATAGTGTCAAACTCAG 301
QY 425 GTCAGATGGTTGGCCAGAGGTGGAAATGAGATAGATCGATCTCTCTGATTTGGAGAA 484
Db 302 GTCAGATGGTTGGCCAGAGGTGGAAATGAGATAGATCGATCTCTCTGATTTGGAGAA 361
QY 485 TGAACATGCCCCCACCACCAAGGTTATGAAGAAGATGCGGCGCATGCCATGATTCGAG 544
Db 362 TGAACATGCCCCCACCACCAAGGTTATGAAGAAGATGCGGCGCATGCCATGATTCGAG 421
QY 545 TTGTGTCCTCATACCAGCGTTCCTCTTTTGTGCTAAAACCCCTGATTTTTCGAAGGAG 604
Db 422 TTGTGTCCTCATACCAGCGTTCCTCTTTTGTGCTAAAACCCCTGATTTTTCGAAGGAG 481
QY 605 CGAATACTACTATTGCTGTATTGTTGGGACCTTTTCGCAATATGAGGAAGAGTGGCAATG 664
Db 482 CGAATACTACTATTGCTGTATTGTTGGGACCTTTTCGCAATATGAGGAAGAGTGGCAATG 541
QY 665 GCATCGTTTACACATGTTGAAAAGACAGTGTGTAATCTCCGAATGCCCAATATAGG 724
Db 542 GCATCGTTTACACATGTTGAAAAGACAGTGTGTAATCTCCGAATGCCCAATATAGG 601
QY 725 TGACCACAGAGAGCGCATGAGTTACTGTGATGGAGTCTTTTAAAGAGGAACCTGGGAGG 784
Db 602 TGACCACAGAGAGCGCATGAGTTACTGTGATGGAGTCTTTTAAAGAGGAAGNCTGGGAGG 661
QY 785 ACAG 788
Db 662 ACAG 665

Search completed: September 12, 2003, 12:46:49
Job time : 422 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	743.4	57.4	57.4	1106	12	US-10-199-673-535	Sequence 535, App
2	743.4	57.4	57.4	1106	12	US-10-187-749-535	Sequence 535, App
3	743.4	57.4	57.4	1106	12	US-10-194-457-535	Sequence 535, App
4	743.4	57.4	57.4	1106	12	US-10-184-644-535	Sequence 535, App
5	743.4	57.4	57.4	1106	12	US-10-196-747-535	Sequence 535, App
6	743.4	57.4	57.4	1106	12	US-10-173-689-535	Sequence 535, App
7	743.4	57.4	57.4	1106	12	US-10-173-690-535	Sequence 535, App
8	743.4	57.4	57.4	1106	12	US-10-173-691-535	Sequence 535, App
9	743.4	57.4	57.4	1106	12	US-10-173-692-535	Sequence 535, App
10	743.4	57.4	57.4	1106	12	US-10-173-693-535	Sequence 535, App
11	743.4	57.4	57.4	1106	12	US-10-173-694-535	Sequence 535, App
12	743.4	57.4	57.4	1106	12	US-10-173-695-535	Sequence 535, App
13	743.4	57.4	57.4	1106	12	US-10-173-699-535	Sequence 535, App
14	743.4	57.4	57.4	1106	12	US-10-173-707-535	Sequence 535, App
15	743.4	57.4	57.4	1106	12	US-10-174-569-535	Sequence 535, App
16	743.4	57.4	57.4	1106	12	US-10-174-587-535	Sequence 535, App


```
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-199-672-535

Query Match      57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTCGGGGAATGTGGGCTGGAGAGGTCTCGCGTGGTACCAAGCTCCAGCCTGCC 103
DB 1 TGGATCTCGGGGAATGTGGGCTGGAGAGGTCTCGCGTGGTACCAAGCTCCAGCCTGCC 60

QY 104 CCAGGACTGCCCTGACCCAGGCGCGCGTCTCGGTGGCAGAGGGCGCGGGAGCG 163
DB 61 CCAGGACTGCCCTGACCCAGGCGCGCGTCTCGGTGGCAGAGGGCGCGGGAGCG 120

QY 164 CCATGGCTGCATCTCGTAAGAGAAAGTCTGTGATGTGTGAGCTTCATAGACGGTTC 223
DB 121 CCATGGCTGCATCTCGTAAGAGAAAGTCTGTGATGTGTGAGCTTCATAGACGGTTC 180

QY 224 TTTTCTCTGGTGTGGCTCTGTAAATGAAGTGAATTTCCCATGTGCTAACTGCT 283
DB 181 TTTTCTCTGGTGTGGCTCTGTAAATGAAGTGAATTTCCCATGTGCTAACTGCT 240

QY 284 TTGGACAACCTGGTACAAAGTGGATFACATTCCTCCATACATACAGGCGGCCCTTCGAA 343
DB 241 TTGGACAACCTGGTACAAAGTGGATFACATTCCTCCATACATACAGGCGGCCCTTCGAA 300

QY 344 CTCACATATGATACATAAATGTGAAGACACAGAGCCCTTTCGAAGTGTGACCTTT 403
DB 301 CTCACATATGATACATAAATGTGAAGACACAGAGCCCTTTCGAAGTGTGACCTTT 360

QY 404 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAATGAGATGATC 463
DB 361 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAATGAGATGATC 420

QY 464 GATCTCTCGATTTGGAGAAATGAACAATGCCCCCAAGGTTATGAAGAAGATGTG 523
DB 421 GATCTCTCGATTTGGAGAAATGAACAATGCCCCCAAGGTTATGAAGAAGATGTG 480

QY 524 GCGCATGACCATGATTCGAGTGTGTCCATACAGCGTCTCTTTGTGTAAAAACC 583
DB 481 GCGCATGACCATGATTCGAGTGTGTCCATACAGCGTCTCTTTGTGTAAAAACC 540

QY 584 CTGATATTTTCAAGGAAGCGAATFACTACTATTTGTGTATTTGGGACCTTTCCGCA 643
DB 541 CTGATATTTTCAAGGAAGCGAATFACTACTATTTGTGTATTTGGGACCTTTCCGCA 600

QY 644 ATATGAGAAAGATGGCAATGGCATCGTTTACACATGTTGAAAGACAGTGTGTA 703
DB 601 ATATGAGAAAGATGGCAATGGCATCGTTTACACATGTTGAAAGACAGTGTGTA 660

QY 704 ATCCGAATGCCCAATATAGTGCACACAGAGAGCGCATGAGTTACTGTGATGAGTT 763
DB 661 ATCCGAATGCCCAATATAGTGCACACAGAGAGCGCATGAGTTACTGTGATGAGTT 720

QY 764 TTAAGAGAAACTGGGAAGACAG 788
DB 721 TTAAGAGAAACTGGGAAGACAG 745
```

RESULT 2
US-10-187-749-535
; Sequence 535, Application US/10187749
; Publication No. US20030153036A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/187,749
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-187-749-535
```

```
Query Match      57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTCGGGGAATGTGGGCTGGAGAGGTCTCGCGTGGTACCAAGCTCCAGCCTGCC 103
DB 1 TGGATCTCGGGGAATGTGGGCTGGAGAGGTCTCGCGTGGTACCAAGCTCCAGCCTGCC 60

QY 104 CCAGGACTGCCCTGACCCAGGCGCGCGTCTCGGTGGCAGAGGGCGCGGGAGCG 163
DB 61 CCAGGACTGCCCTGACCCAGGCGCGCGTCTCGGTGGCAGAGGGCGCGGGAGCG 120

QY 164 CCATGGCTGCATCTCGTAAGAGAAAGTCTGTGATGTGTGAGCTTCATAGACGGTTC 223
DB 121 CCATGGCTGCATCTCGTAAGAGAAAGTCTGTGATGTGTGAGCTTCATAGACGGTTC 180

QY 224 TTTTCTCTGGTGTGGCTCTGTAAATGAAGTGAATTTCCCATGTGCTAACTGCT 283
DB 181 TTTTCTCTGGTGTGGCTCTGTAAATGAAGTGAATTTCCCATGTGCTAACTGCT 240

QY 284 TTGGACAACCTGGTACAAAGTGGATFACATTCCTCCATACATACAGGCGGCCCTTCGAA 343
DB 241 TTGGACAACCTGGTACAAAGTGGATFACATTCCTCCATACATACAGGCGGCCCTTCGAA 300

QY 344 CTCACATATGATACATAAATGTGAAGACACAGAGCCCTTTCGAAGTGTGACCTTT 403
DB 301 CTCACATATGATACATAAATGTGAAGACACAGAGCCCTTTCGAAGTGTGACCTTT 360

QY 404 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAATGAGATGATC 463
DB 361 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCAGAGGTGGGAATGAGATGATC 420
```

Db 361 GTCCATAGTGTCAAACTCAGGTGAGATGGTTGGCCAGAGAGTGGGAAATGAGATAGATC 420
QY 464 GATCCTCCGCAATTTGAGAGATGAACAATGCCCAACCAAGTTATGAAGAAGATGCG 523
Db 421 GATCCCTCCGCAATTTGAGAGATGAACAATGCCCAACCAAGTTATGAAGAAGATGCG 480
QY 524 GCCGCATGACCATGATTCGAGTTGTGTCCCATACACAGCGTTCCCTTTTTCGTAAGAAACC 583
Db 481 GCCGCATGACCATGATTCGAGTTGTGTCCCATACACAGCGTTCCCTTTTTCGTAAGAAACC 540
QY 584 CTGATTATTTTTCAGGAAGCAATGACTACTATTTGTTATTTGGGACCTTTCGCA 643
Db 541 CTGATTATTTTTCAGGAAGCAATGACTACTATTTGTTATTTGGGACCTTTCGCA 600
QY 644 ATATGAGGAAGATGGCAATGGATGTTTACAAACATGTTGAAAGACAGTTGGTATCT 703
Db 601 ATATGAGGAAGATGGCAATGGATGTTTACAAACATGTTGAAAGACAGTTGGTATCT 660
QY 704 ATCCGATGCCCAATATGATGACACACAGAGAACCGCATGAGTTACTGTGATGGAGTTT 763
Db 661 ATCCGATGCCCAATATGATGACACACAGAGAACCGCATGAGTTACTGTGATGGAGTTT 720
QY 764 TTAAGAAGGAAACTGGGAAGGACAG 788
Db 721 TTAAGAAGGAAACTGGGAAGGACAG 745

RESULT 3

US-10-194-457-535
; Sequence 535, Application US/10194457
; Publication No. US20030153037A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RJC296
; CURRENT APPLICATION NUMBER: US/10/194, 457
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA

; ORGANISM: Homo Sapien
US-10-194-457-535
Query Match 57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4,le-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 44 TGGATCTCGGGGAATGCGGCTGGAGAGTCTCCGCTGGTACAGCTCCAGCCTCCGCC 103
Db 1 TGGATCTCGGGGAATGCGGCTGGAGAGTCTCCGCTGGTACAGCTCCAGCCTCCGCC 60
QY 104 CCAGACTGCCCCCTGACCCAGCGGCCCGCTGCTCGGTGGAGAGGCGCGGGAGAG 163
Db 61 CCAGACTGCCCCCTGACCCAGCGGCCCGCTGCTCGGTGGAGAGGCGCGGGAGAG 120
QY 164 CCATGGCCTGCATCCTGAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGGTCC 223
Db 121 CCATGGCCTGCATCCTGAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGGTCC 180
QY 224 TTTTCTCTGCTGGTTGCGCTTGTAAATGAAGTGAATTTCCCATTTGCTACTAACTGCT 283
Db 181 TTTTCTCTGCTGGTTGCGCTTGTAAATGAAGTGAATTTCCCATTTGCTACTAACTGCT 240
QY 284 TTGGACRACCTGGTACAAAGTGGATACCACTTCTCTACATACAGCGCGCCCTTCGAA 343
Db 241 TTGGACRACCTGGTACAAAGTGGATACCACTTCTCTACATACAGCGCGCGCCCTTCGAA 300
QY 344 CTCACATGGATACATAAAATGTGAAGACACAGAGCGCTTTGCAACTGGAGCTGTGACCTTT 403
Db 301 CTCACATGGATACATAAAATGTGAAGACACAGAGCGCTTTGCAACTGGAGCTGTGACCTTT 360
QY 404 GTGCCATAGTGTCAAACTCAGGTGAGTGGTGGCCAGAGAGTGGGAAATGAGATAGATC 463
Db 361 GTGCCATAGTGTCAAACTCAGGTGAGTGGTGGCCAGAGAGTGGGAAATGAGATAGATC 420
QY 464 GATCCTCTGCTATTTGGAGAAATGAACAATGCCCGCCCAAGGTTATGAAGAAGATGTCG 523
Db 421 GATCCTCTGCTATTTGGAGAAATGAACAATGCCCGCCCAAGGTTATGAAGAAGATGTCG 480
QY 524 GCCGCATGACCATGATTCGAGTTGTGTCCCATACACAGCGTTCCCTTTTTCGTAAGAAACC 583
Db 481 GCCGCATGACCATGATTCGAGTTGTGTCCCATACACAGCGTTCCCTTTTTCGTAAGAAACC 540
QY 584 CTGATTATTTTTCAGGAAGCAATGACTACTATTTGTTATTTGGGACCTTTCGCA 643
Db 541 CTGATTATTTTTCAGGAAGCAATGACTACTATTTGTTATTTGGGACCTTTCGCA 600
QY 644 ATATGAGGAAGATGGCAATGGATGTTTACAAACATGTTGAAAGACAGTTGGTATCT 703
Db 601 ATATGAGGAAGATGGCAATGGATGTTTACAAACATGTTGAAAGACAGTTGGTATCT 660
QY 704 ATCCGATGCCCAATATGATGACACACAGAGAACCGCATGAGTTACTGTGATGGAGTTT 763
Db 661 ATCCGATGCCCAATATGATGACACACAGAGAACCGCATGAGTTACTGTGATGGAGTTT 720
QY 764 TTAAGAAGGAAACTGGGAAGGACAG 788
Db 721 TTAAGAAGGAAACTGGGAAGGACAG 745
RESULT 4
US-10-184-642-535
; Sequence 535, Application US/10184642
; Publication No. US20030157635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria

```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C194
; CURRENT APPLICATION NUMBER: US/10/184,642
; CURRENT FILING DATE: 2002-06-27
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-642-535

Query Match          57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTGCGGGAATGTGGCTGGAGAGGTCCTGCGTGGTACAGCTCCAGCCTGCC 103
Db 1 TGGATCTGCGGGAATGTGGCTGGAGAGGTCCTGCGTGGTACAGCTCCAGCCTGCC 60
QY 104 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGGAGCG 163
Db 61 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGGAGCG 120
QY 164 CCATGGCTGCATCTCCTGAAGAAAGTCTGTGATTCGTGAGCTTCATAGCAGCTTC 223
Db 121 CCATGGCTGCATCTCCTGAAGAAAGTCTGTGATTCGTGAGCTTCATAGCAGCTTC 180
QY 224 TTTTCTGCTGGTGTGCTCTTCTTAATGAAGTGAATTCCTCAATGCTAACTGCT 283
Db 181 TTTTCTGCTGGTGTGCTCTTCTTAATGAAGTGAATTCCTCAATGCTAACTGCT 240
QY 284 TTGGACAACTGGTACAAAGTGGATACCATTCCTTACACATACAGCGCCCTTCGAA 343
Db 241 TTGGACAACTGGTACAAAGTGGATACCATTCCTTACACATACAGCGCCCTTCGAA 300
QY 344 CTCACTATGGATACATAAATCTGAAGACACAGAGCTTTGCAACTGGACTGTGACCTTT 403
Db 301 CTCACTATGGATACATAAATCTGAAGACACAGAGCTTTGCAACTGGACTGTGACCTTT 360
QY 404 GTGCCATAGTGTCAAACCTCAGGTGATGTTGGCCAGAGGTTGAGAAATGAGATGATC 463
Db 361 GTGCCATAGTGTCAAACCTCAGGTGATGTTGGCCAGAGGTTGAGAAATGAGATGATC 420
QY 464 GATCTCTCGATTTGGGAATGAAATGCCCAATGCCCAATGCCCAATGCCCAATGCCCA 523
Db 421 GATCTCTCGATTTGGGAATGAAATGCCCAATGCCCAATGCCCAATGCCCAATGCCCA 480
QY 524 GCGCATGACCATGATTCGAGTGTGTGCCATACAGCTTCTCTTTTGTAAAAAACC 583
Db 481 GCGCATGACCATGATTCGAGTGTGTGCCATACAGCTTCTCTTTTGTAAAAAACC 540
QY 584 CTGATATTTTTCAGGAAGCGAATGCTACTATTTGTTGTTGTTGTTGTTGTTGTTGTT 643
Db 541 CTGATATTTTTCAGGAAGCGAATGCTACTATTTGTTGTTGTTGTTGTTGTTGTTGTT 600
QY 644 ATATGAGAAAGATGGCAATGGCATCGTTTACAAATGTTGAAAGACAGTGTGTTATCT 703
Db 601 ATATGAGAAAGATGGCAATGGCATCGTTTACAAATGTTGAAAGACAGTGTGTTATCT 660
QY 704 ATCCGAATGCCCAATATAGTGTACACAGAGAGGCGCATGTTACTGTGATGAGCTTT 763
Db 661 ATCCGAATGCCCAATATAGTGTACACAGAGAGGCGCATGTTACTGTGATGAGCTTT 720
QY 764 TTAAGAGGAAGTGGGAGGACAG 788
Db 721 TTAAGAGGAAGTGGGAGGACAG 745
```

```
RESULT 5
US-10-196-747-535
; Sequence 535, Application US/10196747
; Publication No. US20030162250A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C346
; CURRENT APPLICATION NUMBER: US/10/196,747
; CURRENT FILING DATE: 2002-07-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-196-747-535

Query Match          57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTGCGGGAATGTGGCTGGAGAGGTCCTGCGTGGTACAGCTCCAGCCTGCC 103
Db 1 TGGATCTGCGGGAATGTGGCTGGAGAGGTCCTGCGTGGTACAGCTCCAGCCTGCC 60
QY 104 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGGAGCG 163
Db 61 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTGGTGGAGAGGCGCGGAGCG 120
QY 164 CCATGGCTGCATCTCCTGAAGAAAGTCTGTGATTCGTGAGCTTCATAGCAGCTTC 223
Db 121 CCATGGCTGCATCTCCTGAAGAAAGTCTGTGATTCGTGAGCTTCATAGCAGCTTC 180
QY 224 TTTTCTGCTGGTGTGCTCTTCTTAATGAAGTGAATTCCTCAATGCTAACTGCT 283
Db 181 TTTTCTGCTGGTGTGCTCTTCTTAATGAAGTGAATTCCTCAATGCTAACTGCT 240
QY 284 TTGGACAACTGGTACAAAGTGGATACCATTCCTTACACATACAGCGCCCTTCGAA 343
Db 241 TTGGACAACTGGTACAAAGTGGATACCATTCCTTACACATACAGCGCCCTTCGAA 300
QY 344 CTCACTATGGATACATAAATCTGAAGACACAGAGCTTTGCAACTGGACTGTGACCTTT 403
Db 301 CTCACTATGGATACATAAATCTGAAGACACAGAGCTTTGCAACTGGACTGTGACCTTT 360
QY 404 GTGCCATAGTGTCAAACCTCAGGTGATGTTGGCCAGAGGTTGAGAAATGAGATGATC 463
Db 361 GTGCCATAGTGTCAAACCTCAGGTGATGTTGGCCAGAGGTTGAGAAATGAGATGATC 420
QY 464 GATCTCTCGATTTGGGAATGAAATGCCCAATGCCCAATGCCCAATGCCCAATGCCCA 523
Db 421 GATCTCTCGATTTGGGAATGAAATGCCCAATGCCCAATGCCCAATGCCCAATGCCCA 480
QY 524 GCGCATGACCATGATTCGAGTGTGTGCCATACAGCTTCTCTTTTGTAAAAAACC 583
Db 481 GCGCATGACCATGATTCGAGTGTGTGCCATACAGCTTCTCTTTTGTAAAAAACC 540
QY 584 CTGATATTTTTCAGGAAGCGAATGCTACTATTTGTTGTTGTTGTTGTTGTTGTTGTT 643
Db 541 CTGATATTTTTCAGGAAGCGAATGCTACTATTTGTTGTTGTTGTTGTTGTTGTTGTT 600
```


Db 181 TTTTCCTGCTGGTGTGGCTCTTGTAAATGAAGTGAATTTCCCATGCTACTAACTGCT 240
QY 284 TTGGACAACCTGGTACAAAGTGGATACCAATTCCTTACACATACAGGCGGCCCTTCGAA 343
Db 241 TTGGACAACCTGGTACAAAGTGGATACCAATTCCTTACACATACAGGCGGCCCTTCGAA 300
QY 344 CTCACATGATGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTT 403
Db 301 CTCACATGATGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTT 360
QY 404 GTGCCATGAGTGAACACACAGTCCAGTGTGGCCAGAGGTGGAAATGAGATAGATC 463
Db 361 GTGCCATGAGTGAACACACAGTCCAGTGTGGCCAGAGGTGGAAATGAGATAGATC 420
QY 464 GATCCTCTGCTGATTTGGAGATGAACATGCCGCCACCAAGGTTATGAAGAGATGTGC 523
Db 421 GATCCTCTGCTGATTTGGAGATGAACATGCCGCCACCAAGGTTATGAAGAGATGTGC 480
QY 524 GCCCATGACCATGATTCGAGTTGTGCCATACCAAGCCTTCTCTTTTGTAAATAAAC 583
Db 481 GCCCATGACCATGATTCGAGTTGTGCCATACCAAGCCTTCTCTTTTGTAAATAAAC 540
QY 584 CTGATTAATTTTCAAGAACGGAATACACTACTATTTATTTATTTGGGGACCTTCCGCA 643
Db 541 CTGATTAATTTTCAAGAACGGAATACACTACTATTTATTTATTTGGGGACCTTCCGCA 600
QY 644 ATATGAGGAAGATGGCAATGGCATCGTTTACAAATGTGAAGAGATGTGGATATCT 703
Db 601 ATATGAGGAAGATGGCAATGGCATCGTTTACAAATGTGAAGAGATGTGGATATCT 660
QY 704 ATCCGAATGCCAAATATACGTACCAACAGAGAGCGCATGAGTTACTGTGATGGAGTTT 763
Db 661 ATCCGAATGCCAAATATACGTACCAACAGAGAGCGCATGAGTTACTGTGATGGAGTTT 720
QY 764 TTAAGAGGAAGAACTGGGAGGACAG 788
Db 721 TTAAGAGGAAGAACTGGGAGGACAG 745

RESULT 8

US-10-173-691-535
; Sequence 535, Application US/10173691
; Publication No. US20030166106A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C13
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ. ID NOS: 612
; SEQ. ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-691-535

Query Match 57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4.1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 44 TGGATCTGGGGATGTGGCTGGAGAGGTCCTGGCGTGTACCAAGCTCCAGCCTCCGCC 103

Db 1 TGATCTGGGAATGTGGGCTGGAGAGTCTGCGGTGACAGCCTCCAGCCTGCC 60
QY 104 CCAGGACTGCCCTGACCCAGGCGCGCGCTGCTCGGTGACAGGCGCGCGGAGCG 163
Db 61 CCAGGACTGCCCTGACCCAGGCGCGCGCTGCTCGGTGACAGGCGCGCGGAGCG 120
QY 164 CCATGCGCTGCATCCTGAAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGGCTTCC 223
Db 121 CCATGCGCTGCATCCTGAAGAGAAAGTCTGTGATTCCTGCTGAGCTTCATAGCAGGCTTCC 180
QY 224 TTTTCCTGCTGGTGTGGCTTGTAAATGAAGTGAATTTCCCATGCTACTAACTGCT 283
Db 181 TTTTCCTGCTGGTGTGGCTTGTAAATGAAGTGAATTTCCCATGCTACTAACTGCT 240
QY 284 TTGGACAACCTGGTACAAAGTGGATACCAATTCCTTACACATACAGGCGGCCCTTCGAA 343
Db 241 TTGGACAACCTGGTACAAAGTGGATACCAATTCCTTACACATACAGGCGGCCCTTCGAA 300
QY 344 CTCACATGATGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTT 403
Db 301 CTCACATGATGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGAGCTGTGACCTTT 360
QY 404 GTGCCATGAGTGAACACACAGTCCAGTGTGGCCAGAGGTGGAAATGAGATAGATC 463
Db 361 GTGCCATGAGTGAACACACAGTCCAGTGTGGCCAGAGGTGGAAATGAGATAGATC 420
QY 464 GATCCTCTGCTGATTTGGAGATGAACATGCCGCCACCAAGGTTATGAAGAGATGTGC 523
Db 421 GATCCTCTGCTGATTTGGAGATGAACATGCCGCCACCAAGGTTATGAAGAGATGTGC 480
QY 524 GCCCATGACCATGATTCGAGTTGTGCCATACCAAGCCTTCTCTTTTGTAAATAAAC 583
Db 481 GCCCATGACCATGATTCGAGTTGTGCCATACCAAGCCTTCTCTTTTGTAAATAAAC 540
QY 584 CTGATTAATTTTCAAGAACGGAATACACTACTATTTATTTATTTGGGGACCTTCCGCA 643
Db 541 CTGATTAATTTTCAAGAACGGAATACACTACTATTTATTTATTTGGGGACCTTCCGCA 600
QY 644 ATATGAGGAAGATGGCAATGGCATCGTTTACAAATGTGAAGAGATGTGGATATCT 703
Db 601 ATATGAGGAAGATGGCAATGGCATCGTTTACAAATGTGAAGAGATGTGGATATCT 660
QY 704 ATCCGAATGCCAAATATACGTACCAACAGAGAGCGCATGAGTTACTGTGATGGAGTTT 763
Db 661 ATCCGAATGCCAAATATACGTACCAACAGAGAGCGCATGAGTTACTGTGATGGAGTTT 720
QY 764 TTAAGAGGAAGAACTGGGAGGACAG 788
Db 721 TTAAGAGGAAGAACTGGGAGGACAG 745

RESULT 9

US-10-173-692-535
; Sequence 535, Application US/10173692
; Publication No. US20030166188A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C20
; CURRENT FILING DATE: 2002-06-17
; CURRENT APPLICATION NUMBER: US/10173,692
; CURRENT FILING DATE: 2002-06-17

```

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-692-535

```

Query Match	57.4%;	Score 743.4;	DB 12;	Length 1106;	
Best Local Similarity	99.9%;	Prod. No. 4.1e-237;			
Matches 744;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	44	TGGATCTGCGGGAAATGTGGCGTGGAGAGTCCCTGCGTGGTACACAGAGCTCCACAGCTGGCC	103		
Db	1	TGGATCTGCGGGAAATGTGGCGTGGAGAGTCCCTGCGTGGTACACAGCTCCACAGCTGGCC	60		
QY	104	CCAGGACTGCCCTGACCCAGCGCGCCGCTGCTGGTGGCAGAGAGCGCGCGGAGG	163		
Db	61	CCAGGACTGCCCTGACCCAGCGCGCCGCTGCTGGTGGCAGAGAGCGCGCGGAGG	120		
QY	164	CCATGGCCCTGCATCTGAAGAGAAAGTCTGTGATGTCTGTGAGCTTCATAGCAGCTGCC	223		
Db	121	CCATGGCCCTGCATCTGAAGAGAAAGTCTGTGATGTCTGTGAGCTTCATAGCAGCTGCC	180		
QY	224	TTTTTCCTGCTGGTGTGGCGCTTGTAAATGAAGTGAATTCCTCATTTGCCATTGCTAACTGCT	283		
Db	181	TTTTTCCTGCTGGTGTGGCGCTTGTAAATGAAGTGAATTCCTCATTTGCCATTGCTAACTGCT	240		
QY	284	TTGGACAACCTGGTACAAAGTGGATACCATTCCTTACACATPACAGGCGGCCCTTTCGAA	343		
Db	241	TTGGACAACCTGGTACAAAGTGGATACCATTCCTTACACATPACAGGCGGCCCTTTCGAA	300		
QY	344	CTCACTATGGATACATAAATGTGAAGACACAGAGCGCTTGCACATGSACTGTGACCTTT	403		
Db	301	CTCACTATGGATACATAAATGTGAAGACACAGAGCGCTTGCACATGSACTGTGACCTTT	360		
QY	404	GTGCCATAGTGTCAAACTCAGCTCAGATGGTGGCCAGAGGTGGGAAATGAGATAGATC	463		
Db	361	GTGCCATAGTGTCAAACTCAGCTCAGATGGTGGCCAGAGGTGGGAAATGAGATAGATC	420		
QY	464	GATCCTCCTGCATTTGGAGATGAACAAATGCCCCACCAAGGTTATGAAGAGATGTGC	523		
Db	421	GATCCTCCTGCATTTGGAGATGAACAAATGCCCCACCAAGGTTATGAAGAGATGTGC	480		
QY	524	GCCCATGACCATGATTCGAGTTGTGTGCCATACACAGCGTTCCTCTTTTGCCTAAAAAC	583		
Db	481	GCCCATGACCATGATTCGAGTTGTGTGCCATACACAGCGTTCCTCTTTTGCCTAAAAAC	540		
QY	584	CTGATATATTTTTCAGGAGCGAATACTACTATTTTGTGTTATTTGGGGACCTTTCCGCA	643		
Db	541	CTGATATATTTTTCAGGAGCGAATACTACTATTTTGTGTTATTTGGGGACCTTTCCGCA	600		
QY	644	ATATAGGAAAGATGGCAATGGCATCGCTTTACACATGTTGAAGAGACAGATTTGGTATCT	703		
Db	601	ATATAGGAAAGATGGCAATGGCATCGCTTTACACATGTTGAAGAGACAGATTTGGTATCT	660		
QY	704	ATCCGAATGCCAAATATACGTGACCACAGAGAGCGCATGAGTTACTGTGATGGAGTTT	763		
Db	661	ATCCGAATGCCAAATATACGTGACCACAGAGAGCGCATGAGTTACTGTGATGGAGTTT	720		
QY	764	TTAAGAGAAACTGGGAAGACAG	788		
Db	721	TTAAGAGAAACTGGGAAGACAG	745		

```

RESULT 10
US-10-173-694-535
; Sequence 535, Application US/10173694
; Publication No. US20030166107A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc

```

```

? APPLICANT: Goddard,Audrey
? APPLICANT: Godowski,Paul J.
? APPLICANT: Gurney,Austin L.
? APPLICANT: Pan,James
? APPLICANT: Smith,Victoria
? APPLICANT: Watanabe,Colin K.
? APPLICANT: Wood,William I.
? APPLICANT: Zhang,Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? TYPE OF INVENTION: ACIDS ENCODING THE SAME
? FILE REFERENCE: P3430R1C19
? CURRENT APPLICATION NUMBER: US/10/173,694
? CURRENT FILING DATE: 2002-06-17
? Prior Application removed - See File Wrapper or Palm
? NUMBER OF SEQ ID NOS: 612
? SEQ ID NO 535
? LENGTH: 1106
? TYPE: DNA
? ORGANISM: Homo Sapien
? US-10-173-694-535

```

Query Match	57.4%;	Score	743.4;	DB	12;	Length	1106;
Best Local Similarity	99.9%;	Prod. No.	4.1e-237;				
Matches	744;	Conservative	0;	Mismatches	1;	Indels	0;
Qy	44	TGGATCTGCGGGAATCTGGCGTGCAGAGGTCCTGCGGTGTACAGAGCTCCAGAGCTGCC	103				
Db	1	TGGATCTGCGGGAATCTGGCGTGCAGAGGTCCTGCGGTGTACAGAGCTCCAGAGCTGCC	60				
Qy	104	CCAGGACTGCCCTGACCCAGAGCGCGCCGCTGCTCGGTGGCAGAGGCGCGCGAGAGCG	163				
Db	61	CCAGGACTGCCCTGACCCAGAGCGCGCCGCTGCTCGGTGGCAGAGGCGCGCGAGAGCG	120				
Qy	164	CCATGGGCTGCATCTGGAAGAGAAAGTCTCTGATTCGTGTGAGCTTCATAGCAGAGTTCC	223				
Db	121	CCATGGGCTGCATCTGGAAGAGAAAGTCTCTGATTCGTGTGAGCTTCATAGCAGAGTTCC	180				
Qy	224	TTTTCTGCTGGTGTGGGCTCTCTGAATGAAGTGAATTTCCCACTTCGACTAAACTGCT	283				
Db	181	TTTTCTGCTGGTGTGGGCTCTCTGAATGAAGTGAATTTCCCACTTCGACTAAACTGCT	240				
Qy	284	TGGACAACTCGTACAAAGTGGATACCATTCCTTACACATACAGGCGGCCCTTCGAA	343				
Db	241	TGGACAACTCGTACAAAGTGGATACCATTCCTTACACATACAGGCGGCCCTTCGAA	300				
Qy	344	CTCACTATGATACATATAATGTGAAGACACAGAGCGCTTTGCAACTGGACTGTGACCTTT	403				
Db	301	CTCACTATGATACATATAATGTGAAGACACAGAGCGCTTTGCAACTGGACTGTGACCTTT	360				
Qy	404	GTGCCATAGTGTCAAACTCAGGTCAGATGTTGGCCAGAAAGTGGGAAATGAGATAGATC	463				
Db	361	GTGCCATAGTGTCAAACTCAGGTCAGATGTTGGCCAGAAAGTGGGAAATGAGATAGATC	420				
Qy	464	GATCCTCCTGCATTTGGAGAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTGC	523				
Db	421	GATCCTCCTGCATTTGGAGAATGAACAATGCCCCACCAAGGTTATGAAGAGATGTGC	480				
Qy	524	GCCGCATGACCATGATTCGAGTTGTGTCCCATACAGAGGTTCTCTTTGCTTAAAAAACC	583				
Db	481	GCCGCATGACCATGATTCGAGTTGTGTCCCATACAGAGGTTCTCTTTGCTTAAAAAACC	540				
Qy	584	CTGATATTTTTTCAAGGAAGCAATACTACTATTGTGTTATTTGGGACCTTTCCCGCA	643				
Db	541	CTGATATTTTTTCAAGGAAGCAATACTACTATTGTGTTATTTGGGACCTTTCCCGCA	600				
Qy	644	ATATGAGGAAGATGGCAATGSGCATCGTTTACAACATCTTGAAAAAGACAGTTTGTGTCT	703				
Db	601	ATATGAGGAAGATGGCAATGSGCATCGTTTACAACATCTTGAAAAAGACAGTTTGTGTCT	660				
Qy	704	ATCCGAATGCCAAATATATACGTGACACAGAGAGCGGATGAGTTACTGTGATGAGATTT	763				
Db	661	ATCCGAATGCCAAATATATACGTGACACAGAGAGCGGATGAGTTACTGTGATGAGATTT	720				


```
Db 301 CTCACATGATGATACATAAATGTGAAGACACAAAGAGCCCTTTGCAACTGGAGTGTGACCTTT 360
QY 404 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCCAAGAGTGGGAAATGAGATAGATC 463
Db 361 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCCAAGAGTGGGAAATGAGATAGATC 420
QY 464 GATCCTCCTGCTGATTTGGAGAAATGAACAATGCCGCCCAACAAAGTTATGAAGAAGATGTG 523
Db 421 GATCCTCCTGCTGATTTGGAGAAATGAACAATGCCGCCCAACAAAGTTATGAAGAAGATGTG 480
QY 524 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
Db 481 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 584 CTGATTTATTTTCAAGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 643
Db 643 CTGATTTATTTTCAAGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 644 ATATGAGAAAGATGGCAATGGATCGTTTACCAATGTTGAAAGACAGATGGTATCT 703
Db 703 ATATGAGAAAGATGGCAATGGATCGTTTACCAATGTTGAAAGACAGATGGTATCT 660
QY 704 ATCCGATGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
Db 763 ATCCGATGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 764 TTAAGGAAGCAATGGGAAGGACAG 788
Db 788 TTAAGGAAGCAATGGGAAGGACAG 745

RESULT 13
US-10-173-707-535
; Sequence 535, Application US/10173707
; Publication No. US20030166110A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desborders, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C17
; CURRENT APPLICATION NUMBER: US/10/173,707
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; TYPE: DNA
; LENGTH: 1106
; ORGANISM: Homo Sapien
US-10-173-707-535

Query Match 57.4%; Score 743.4; DB 12; Length 1106;
Best Local Similarity 99.9%; Pred. No. 4,1e-237;
Matches 744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 TGGATCTGCGGATGTTGGGTGGAGAGTCCGCGGTACAGCTCCAGCTGCCCTGCC 103
Db 1 TGGATCTGCGGATGTTGGGTGGAGAGTCCGCGGTACAGCTCCAGCTGCCCTGCC 60
QY 104 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTCGTGGCAGAGGCGCGCGGAGG 163
Db 61 CCAGGACTGCCCTGACCCAGCGCGCGCTGCTCGTGGCAGAGGCGCGCGGAGG 120
QY 164 CCATGGCTGCATCTGATGAGAAAGTCTGTGATTCCTGTGAGCTTATAGCAGCGTTC 223
```

```
Db 121 CCATGGCTGCATCTGATGAGAAAGTCTGTGATTCCTGTGAGCTTCATAGCAGCGTTC 180
QY 224 TTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 283
Db 181 TTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 284 TTGGACAAACCTGGTACAAAAGTGGATACCATCTCCCTACACATACAGGCGGCCCTTCGAA 343
Db 241 TTGGACAAACCTGGTACAAAAGTGGATACCATCTCCCTACACATACAGGCGGCCCTTCGAA 300
QY 344 CTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
Db 301 CTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
QY 404 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCCAAGAGTGGGAAATGAGATAGATC 463
Db 463 GTGCCATAGTGTCAAACTCAGGTCAGATGGTTGGCCCAAGAGTGGGAAATGAGATAGATC 420
QY 464 GATCCTCCTGCTGATTTGGAGAAATGAACAATGCCGCCCAACAAAGTTATGAAGAAGATGTG 523
Db 421 GATCCTCCTGCTGATTTGGAGAAATGAACAATGCCGCCCAACAAAGTTATGAAGAAGATGTG 480
QY 524 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
Db 481 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 584 CTGATTTATTTTCAAGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 643
Db 643 CTGATTTATTTTCAAGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 644 ATATGAGAAAGATGGCAATGGATCGTTTACCAATGTTGAAAGACAGATGGTATCT 703
Db 703 ATATGAGAAAGATGGCAATGGATCGTTTACCAATGTTGAAAGACAGATGGTATCT 660
QY 704 ATCCGATGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
Db 763 ATCCGATGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 764 TTAAGGAAGCAATGGGAAGGACAG 788
Db 788 TTAAGGAAGCAATGGGAAGGACAG 745

RESULT 14
US-10-174-569-535
; Sequence 535, Application US/10174569
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desborders, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C39
; CURRENT APPLICATION NUMBER: US/10/174,569
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 535
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-569-535
```


Search completed: September 12, 2003, 16:29:44
Job time : 366 secs

```

RESULT 15
US-10-174-583-535
; Sequence 535, Application US/10174583
; Publication No. US20030166112A1
; GENERAL INFORMATION:
; APPLICANT: Baker Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli

```


High quality sequence stop: 523.

```
FEATURES
  Location/Qualifiers
    1..794
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:30341233"
      /tissue_type="Human Placenta"
      /lab_host="PH10B Tona"
      /clone_lib="NIH_MGC_147"
      /notes="Organ: placenta; Vector: pBluescriptR; Site_1:
      all-101; Site_2: BamH; Oligo-dT primed using primer
      5'-TTTTTTTTTTTTTNN-3', size-selected for average
      insert size 2.3 kb and normalized to 80T 5. This is a
      primary library enriched for full-length clones and
      constructed using the Cap-trapper method (Carninci, in
      preparation). Library constructed by M. Brownstein
      (NIMH/NHGRI, National Institutes of Health). Note: This is
      a NIH_MGC library."
BASE COUNT      204 a      185 c      205 g      200 t
ORIGIN
Query Match      49.9%; Score 646.2; DB 14; Length 794;
Best Local Similarity 99.4%; Pred. No. 1.3e-180;
Matches 659; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 28 CCGCGGTCCTTATTGGATCTCGGGATGTGGCTGGAGAGCTCGCGGTGACC 87
Db 24 CAGGGTCCTTATTGGATCTCGGGATGTGGCTGGAGAGCTCGCGGTGACC 83
QY 88 AGCTCCAGGCTCCCGCCAGGACTCCCGTGAACAGGCGCGCGCTCGGTGGCAG 147
Db 84 AGCTCCAGGCTCCCGCCAGGACTCCCGTGAACAGGCGCGCGCTCGGTGGCAG 143
QY 148 GAGGCGCGGGAGCGCCATGCGCTGACCTGAGAGAAAGTCTGATGCTGTGAGC 207
Db 144 GAGGCGCGGGAGCGCCATGCGCTGACCTGAGAGAAAGTCTGATGCTGTGAGC 203
QY 208 TTCTAGACAGCTTCTTTCTCTGCTGCTGTGCTGTGTAAGAAAGTGAATTCCTCA 267
Db 204 TTCTAGACAGCTTCTTTCTCTGCTGCTGTGCTGTGTAAGAAAGTGAATTCCTCA 263
QY 268 TTGCTACTAAAGTCTTTGGCAACCTGGTACAAAGTGGATACCAATTCCTCACATAC 327
Db 264 TTGCTACTAAAGTCTTTGGCAACCTGGTACAAAGTGGATACCAATTCCTCACATAC 323
QY 328 AGCGGCGCCCTTGGAACTCACTATGATACATAAATGTGAAGACACAGAGCCTTGCCAA 387
Db 324 AGCGGCGCCCTTGGAACTCACTATGATACATAAATGTGAAGACACAGAGCCTTGCCAA 383
QY 388 CTGGAGCTGACCTTTGTGCCATAGTGTCAAACTCAGGTGAGATGTTGGCCAGAGGTG 447
Db 384 CTGGAGCTGACCTTTGTGCCATAGTGTCAAACTCAGGTGAGATGTTGGCCAGAGGTG 443
QY 448 GGAATAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
Db 444 GGAATAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 503
QY 508 TATGAAGAAGATGTCGGCGGCGATGACCATGATGATGATGATGATGATGATGATGATG 567
Db 504 TATGAAGAAGATGTCGGCGGCGATGACCATGATGATGATGATGATGATGATGATGATG 563
QY 568 CTTTGTGCT-AAAAACCTGATATTTTTCAGGAAGCGAATGATGATGATGATGATGATG 626
Db 564 CTTTGTGCTAAAAACCTGATATTTTTCAGGAAGCGAATGATGATGATGATGATGATG 623
QY 627 TTGGGGACCTTTCGCAATATGAGGAAGATGCGCAATGCGCAATGCGCAATGCGCAATG 686
Db 624 TTGGGGACCTTTCGCAATATGAGGAAGATGCGCAATGCGCAATGCGCAATGCGCAATG 683
QY 687 AAA 689
Db 684 AAA 686
```

```
RESULT 2
BX280331
LOCUS
DEFINITION
  IMAGE:5724478, mRNA sequence.
ACCESSION
  BX280331
VERSION
  BX280331.1 GI:28612369
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 505)
  Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
  Radelof,U., Schneider,D. and Korn,B.
  Human Unigeneset - RZPD3
  Unpublished
  Contact: Ina Rofls
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
  RZPD; IMAGp998M2312713.
  RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
  Human Unigeneset - RZPD3 (RZPDLIB No.972)
  http://www.rzpd.de/CloneCards/cgi-
  bin/showlib.pl.cgi?response?libNo=972 Contact: Ina Rofls
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Heubnerweg 6, D-14059 Berlin, Germany
  Tel: +49 30 32639 101
  Fax: +49 30 32639 111
  www.rzpd.de
  This clone is available royalty-free from RZPD;
  contact RZPD (clone.rzpd.de) for further information. Seq primer:
  M13u, Primer sequence: CGTGTAAACGACGCGCCAGT.
FEATURES
  Location/Qualifiers
    1..505
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGp998M2312713 ; IMAGE:5724478"
      /clone_lib="NIH_MGC_125"
      /lab_host="PH10B"
      /notes="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
      Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
      of three ovaries, from females ranging in age from 38 to
      49 yo. Library is oligo-dT primed and directionally cloned
      (EcoRV site is destroyed upon cloning). Average insert
      size 2.1 kb, insert size range 1-3.5 kb. Library is
      normalized and enriched for full-length clones and was
      constructed by C. Gruber (Invitrogen). Research Genetics
      tracking code 036."
BASE COUNT      154 a      103 c      119 g      129 t
ORIGIN
Query Match      38.3%; Score 495.4; DB 13; Length 505;
Best Local Similarity 99.8%; Pred. No. 8.4e-136;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 292 CCGTGTCAAAAGTGGATACCATCTCTACACATACAGGCGCGCCCTTCGAACTCATAT 351
Db 1 CCGTGTCAAAAGTGGATACCATCTCTACACATACAGGCGCGCCCTTCGAACTCATAT 60
QY 352 GGTATACATAAATGTGAAGACACAGAGCCCTTGCACACTGGACTGTGACCTTTGTGCCATA 411
Db 61 GGTATACATAAATGTGAAGACACAGAGCCCTTGCACACTGGACTGTGACCTTTGTGCCATA 120
QY 412 GTGTCAAACTCAGGTGATGTTGGCCAGAGTGGGAAATGAGATGATGATGATGATGATG 471
Db 121 GTGTCAAACTCAGGTGATGTTGGCCAGAGTGGGAAATGAGATGATGATGATGATGATG 180
QY 472 TGCATTTGGAGATGAACATGCCCCCACCAGAGTTATGAGAGAGATGTCGGCGGCGATG 531
```



```
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 650V"
/notes="vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
BASE COUNT      176 a   166 c   166 g   192 t
ORIGIN
Query Match      37.7%; Score 487.6; DB 14; Length 700;
Best Local Similarity 91.3%; Pred. No. 2e-133;
Matches 517; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 64 CTGGAGAGGTCCTCCGCTGGTACCAAGCCTCCAGCCTCCGCCCGAGGACTGCCCTGACGCCA 123
Db 1 CTGGAGGCGTCTCGAGCGGTACCAAGCCTCCAGCCTCCGCCCGAGGACTGCCCTGACGCCA 60
QY 124 GCGCGCCGCTGCTCGGTGGCAGGAGGCGCGGAGCGCCATGCCCTGCATCCTGAAG 183
Db 61 GCGCGCCGCTGCTCGGTGGCAGGAGGCGCGGAGCGCCATGCCCTGCATCCTGAAG 120
QY 184 AGAAGTCTGTGATGCTGTGAGCTTCATPAGCAGCTTCCTTTTCTGCTGGTGTGGT 243
Db 121 AGAAGTCTGTGATGCTGTGAGCTTCATPAGCAGCTTCCTTTTCTGCTGGTGTGGT 180
QY 244 CTGTGAATGAAGTGAATTCCTCATCTGCTTAACTGCTTTGGACAACCTGGTACAAAG 303
Db 181 CTGTGAATGAAGTGAATTCCTCATCTGCTTAACTGCTTTGGACAACCTGGTACAAAG 240
QY 304 TGGATACCATCTCTCTACATACAGCGGCGCCCTTCGAACTCACATGATGATACATAAT 363
Db 241 TGGATACCATCTCTCTATACATACAGCGGCGCCCTTCGAACTCACATGATGATACATAAT 300
QY 364 GTGAGACACAGAGCCTTTGCACTGCACTGTGACCTTTGCGCATAGTGTCAACTCA 423
Db 301 GTGAGAGCGCAAGAGCCTTTGCACTGCACTGTGACCTTTGCGCATAGTGTCAACTCG 360
QY 424 GGTCAAGTGGTGGCCAGAGGTGGGAAATGAGATAGATCGATCCCTGCAATTTGGAGA 483
Db 361 GGTCAAGTGGTGGCCAGAGGTGGGAAATGAGATAGATCGATCCCTGCAATTTGGAGA 420
QY 484 ATGAACAATGCCGCCACCAAGGTTATGAAGAAGATGCGGCCGCGATGACCATGATTCGA 543
Db 421 ATGAACAATGCAACCCACCAAGGCTATGAAGAGGATGTTGGCGCATGACATGATTCG 480
QY 544 GTTGTGTCATACACAGCTTCCTCTTTGCTTAAACCCCTGATTAATTTTTCGAAGAA 603
Db 481 GTCTGATCCATACAGCGTCTCTTTTACTGAAACCCCTGATTAATTTTTCGAAGAA 540
QY 604 GCGAATACTACTATTGTGTATTG 629
Db 541 GCGAAGCATATTTTACATCAATTG 566
```

```
RESULT 5
BM547511
LOCUS
DEFINITION BM547511 1088 bp mRNA linear EST 20-FEB-2002
AGENCOURT 6507757 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724478
5', mRNA sequence.
BM547511
ACCSSION BM547511.1 GI:18781340
VERSION BM547511
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1088)
NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
```

```
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12713 row: m column: 23
High quality sequence stop: 682.
FEATURES
Location/Qualifiers
source
1..1088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5724478"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/notes="Orgn: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."
BASE COUNT      325 a   205 c   219 g   269 t   70 others
ORIGIN
Query Match      35.9%; Score 454.4; DB 12; Length 1088;
Best Local Similarity 99.8%; Pred. No. 2e-126;
Matches 465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 323 CATACAGCGGCCCTTCGAACTCACTATGGATACATAAATGTGAAGACACAAAGCCTT 382
Db 29 CATACAGCGGCCCTTCGAACTCACTATGGATACATAAATGTGAAGACACAAAGCCTT 88
QY 383 TCCAACTGGAGTGTGACCTTTGTGCCATAGTGTCAAACTCAGGTGAGTGTGCCAGA 442
Db 89 TCCAACTGGAGTGTGACCTTTGTGCCATAGTGTCAAACTCAGGTGAGTGTGCCAGA 148
QY 443 AGGTGGGAAATGAGATAGATCGATCCTCTGATTTGGAGAATGAACAATGCCGCCACA 502
Db 149 AGGTGGGAAATGAGATAGATCGATCCTCTGATTTGGAGAATGAACAATGCCGCCACA 208
QY 503 AAGGTTATGAAGAAGATGTGCGCCGATGACCATGATTCGAGTTGTGCCATACCCAGC 562
Db 209 AAGGTTATGAAGAAGATGTGCGCCGATGACCATGATTCGAGTTGTGCCATACCCAGC 268
QY 563 TTCCCTCTTTTCTAAAAAACCTGATTTATTTTCAAGGAAGCAATACTACTATTGTG 622
Db 269 TTCCCTCTTTTCTAAAAAACCTGATTTATTTTCAAGGAAGCAATACTACTATTATG 328
QY 623 TTATTGGGGACCTTCCGCAATATGAGAAAGATGGCAATGGCATGCGTTTACACATGT 682
Db 329 TTATTGGGGACCTTCCGCAATATGAGAAAGATGGCAATGGCATGCGTTTACACATGT 388
QY 683 TGAAGAAACAGAGTTGGTATCTATCCGAATGCCAAATATACGTGACACAGAGAGCGCA 742
Db 389 TGAAGAAACAGAGTTGGTATCTATCCGAATGCCAAATATACGTGACACAGAGAGCGCA 448
QY 743 TGAGTTACTGTGATGGAGTTTTTAAAGAGAACTGGGAAGGACAG 788
Db 449 TGAGTTACTGTGATGGAGTTTTTAAAGAGAACTGGGAAGGACAG 494
RESULT 6
CA318140
LOCUS
DEFINITION CA318140 766 bp mRNA linear EST 26-NOV-2002
UI-M-FW0-cbg-p-16-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
IMAGE: 6813281 5', mRNA sequence.
CA318140
ACCSSION CA318140
VERSION CA318140.1 GI:24536264
```

```

KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus

REFERENCE
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)

FEATURES
    source
        Location/Qualifiers
            1..766
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /clone="IMAGE: 6813281"
            /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="NIH_BMAP_FW0"
            /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
            Site_2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            primer containing a Not I site. Double strand cDNA was
            size selected according to mRNA size fraction, ligated
            with EcoR I adaptor, digested with NotI and then cloned
            directionally into pYX-Asc vector. The library tag
            sequence located between the Not I site and the polyA tail
            is ACCGAGACAG. This library was created for the University
            of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
            Developing Mouse Nervous System', supported by National
            Institute of Mental Health (NIMH), Hemm Chin, Ph.D.,
            Program coordinator."
            201 a 191 c 189 g 184 t 1 others

BASE COUNT
ORIGIN
    Query Match      34.7%; Score 449.4; DB 14; Length 766;
    Best local Similarity 81.7%; Pred. No. 5e-122;
    Matches 519; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 160 AGCGCATGGCTGATCCTGAGAGAAAGCTGTGATGTGTGAGCTTCATAGCAGCG 219
    |||||
Dd 27 AGCGCATGGCTGATCCTGAGAGAAAGCTGTGATGTGTGAGCTTCATAGCAGCG 86
    |||||
QY 220 TTCCTTTTCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 279
    |||||
Dd 87 TGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146
    |||||
QY 280 TGCCTTGGACAACTGGTACAAAGTGGATACCACTTCTCTACATACAGAGCGGCCCTT 339
    |||||
Dd 147 TGCCTTGGACAACTGGTACAAAGTGGATACCACTTCTCTACATACAGAGCGGCCCTT 206
    |||||
QY 340 CGACTCCTATGATGATATATATATATATATATATATATATATATATATATATATAT 399
    |||||
Dd 207 CGACTCCTATGATGATATATATATATATATATATATATATATATATATATATATAT 266
    |||||
QY 400 CTTTGTGCCATAGTGTCAAACTCAGGTTCAGTGTGTCGACAGAGAGTGGGAAATGAGATA 459
    |||||
Dd 267 CACTGTGCCATAGTGTCAAACTCAGGTTCAGTGTGTCGACAGAGAGTGGGAAATGAGATA 326

```

```

460 GATCGATCCTCTGCAATTTGGAGAAATGAACATGCCCCCAACCAAGTTATGAAAGAT 519
    |||
Dd 327 GACCATGATCCTCTGCAATTTGGAGAAATGAACATGCCCCCAACCAAGTTATGAAAGAT 386
    |||||
QY 520 GTGCGCGCATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 579
    |||||
Dd 387 GTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
    |||||
QY 580 AACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
    |||||
Dd 447 ATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
    |||||
QY 640 CGCATATGAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
    |||||
Dd 507 CGCATATGAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
    |||||
QY 700 ATCTATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
    |||||
Dd 567 GCCTATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
    |||||
QY 760 GTTTTAAAGAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
    |||||
Dd 627 GTTTTAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 686
    |||||

RESULT 7
CA319843
LOCUS      791 bp      mRNA      linear      EST 26-NOV-2002
DEFINITION      UI-M-FW0-cca-1-12-0-01.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
IMAGE: 6817021 5', mRNA sequence.
CA319843
ACCESSION      CA319843.1 GI:24537967
VERSION        CA319843.1
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Mus musculus
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/
AUTHORS        National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE          Unpublished
JOURNAL        Contact: Robert Strausberg, Ph.D.
COMMENT        Email: cgabbs-r@mail.nih.gov
                Tissue Procurement: Dr. Jim Lin, University of Iowa
                cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                This clone was contributed by the Brain Molecular Anatomy Project
                (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
    1..791
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6"
    /db_xref="taxon:10090"
    /clone="IMAGE: 6817021"
    /tissue_type="whole brain"
    /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
    /lab_host="DH10B (T1 phage resistant)"
    /clone_lib="NIH_BMAP_FW0"
    /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
    Site_2: Not I; The library was constructed according
    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. Denatured RNA was size fractionated on a 1% agarose
    gel. First strand cDNA synthesis was primed with oligo-dT
    primer containing a Not I site. Double strand cDNA was
    size selected according to mRNA size fraction, ligated
    with EcoR I adaptor, digested with NotI and then cloned

```


Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>; Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: <http://fantom.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

COMMENT

FEATURES
source

1. 2927
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9430041E16"
/db_xref="taxon:10090"
/clone="9430041E16"
/tissue_type="embryonic body between diaphragm region and neck"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
74. .991
/notes="unnamed protein product; putative sialyltransferase 7 ((alpha-N-acetylneuraminyl 2-3-beta-galactosyl-1,3)-N-acetyl galactosaminide alpha-2,6-sialyltransferase) C (MGDI:1341828, GBINL_011372, evidence: BLASTN, 99%, match=2923)"
/codon_start=1
/protein_id="BAC28836.1"
/db_xref="GI:26330212"

CDS

RESULT 9
BB626911
LOCUS
DEFINITION
BB626911 RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck Mus musculus cDNA clone
9430041E16 5', mRNA sequence.

ACCESSION

BB626911
VERSION
BB626911.1 GI:15398967

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 659)

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,Y., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished

TITLE

Yoshihide Hayashizaki

JOURNAL

Unpublished

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
M., Konno,H., Okazaki,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)

BASE COUNT
ORIGIN

Query Match 34.7%; Score 449.4; DB 11; Length 2927;
Best Local Similarity 81.7%; Pred. No. 8.7e-122;
Matches 519; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Qy 160 AGCGCATGCGTCGATCCTGAAGAGAAAGTCGTGATGCTGTGAGCTTCATAGCAGCG 219
|||||
Db 68 AGCGCATGCGTCGATCCTGAAGAGAAAGTCGTGATGCTGTGAGCTTCATAGCAGCG 127
Qy 220 TTCCTTTCTGCTGTGCTGTGCTGTCTGTAATGAAGTGAATTCCTACTACTAAAC 279
|||||
Db 128 TGCACTGTGCTGCGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
Qy 280 TGCTTTGACAACTGGTCAAAAGTGGATACCATTCCTCTACACATACAGCGCGCCCTT 339
|||||
Db 188 TGCTTTGACAGCGCTAAGACCAATGGATCCCTTTGCCCTTACACATTCAGGCGCCCTT 247
Qy 340 CGAATCTACTATGATACATAAATGTGAGACACAGAGCGCTTCGCACTGGACTGAC 399
|||||
Db 248 CGAATCTACTATGATACATAAATGTGAGACACAGAGCGCTTCGCACTGGACTGAC 307
Qy 400 CTTTGTGCAATGTCGCAACTGAGTCAGATGCTGGCCAGAGGTGGAAATACAGATA 459
|||||
Db 308 CACTGTGCAATGTCGCAACTGAGTCAGATGCTGGCCAGAGGTGGGGAGAGATA 367
Qy 460 GATCGATCCTCCTGCATTTTGGAGAAATGAACAATGCCCCACCACCAAGGTTATGAAGAAG 519


```

/dev_stage="13 days embryo"
/lab_host="DHI0B"
/clone_lib="RIKEN full-length enriched, 13 days embryo
heart"
/notes="Site.1: SalI; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in RIKEN
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCGCCAACTGTTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGTCGAGATTAATTAATTAATCCCGCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
phluescript KS(-) after bulk excision from Lambda FLC I."
158 a 158 c 167 g 152 t

```

Query Match	30.0%;	Score 388.8;	DB 10;	Length 635;
Best Local Similarity	81.7%;	Prod. No. 5.1e-104;		
Matches 461;	Conservative	0;	Mismatches 102;	Indels 1; Gaps 1;
QY	160	AGGCCCATGGCGTCATCCTGAAGAGAAAAGTCTGTGATTCGTGTGAGCTTCATAGCAGCG	219	
Db	73	AGGCCATGGCGTCGATCCTCAAGAGAAAGCCCGTCGTGTGGTGAGCTTCATAGCCCTG	132	
QY	220	TTCCTTTTTCCTGCTGGTTGTCGTCCTGTGTAATGAAGTGAATTTCCCATTTGCTACTAATAC	279	
Db	133	TGCATCTTCTGCTAGCCATGCGCCTTGTCAATGATGCGACTTTCCTTTGCTCTCGAAC	192	
QY	280	TGCTTTGGACAACCTGGTGTACAAATGGATACCAATTCCTCTACACATACAGGCGGCCCTT	339	
Db	193	TGCTTTGGACAGCCCTAAGACAATAATGGATCCCTTTGCGCTACACATTCAGGCGAGCTCTT	252	
QY	340	CGAACTCACTATGGATACATAAAATGTGAAGACACAAGAGCCCTTGTCAACTGGACATGTGCAC	399	
Db	253	CGAACTCACTATGGATACATAAATGTGAGGACCCAAAGAGCCCTTGTCAACTGACTGTATAC	312	
QY	400	CTTTTGGCCATAGTGTCAAACCTCAGGTCAGATGGTTGGCCAGAGAGTGGGAATCAGATA	459	
Db	313	CACGTGGCCATCGTGTCAAAATTCAGGTCAAAATGGTTCGGGCGAGAGTGGGGGAAGAGATT	372	
QY	460	GATCGATCCCTCGCATTTGGAGAAATGAACAATGCCGCCACAAAGGTTATGAAGAAGAT	519	
Db	373	GACCATGCATCTGTCATCTGGAGAAATGAACAAGCGCCCAACAAAGGCTTTGAGGAAGAT	432	
QY	520	GTGGGCGCATGACCATGATTCGAGTTGTGTCCCATACCCAGCGTCTCTCTTTGCTABAA	579	
Db	433	GTGGGCTACATGACAAATGGTCCGAGTTGTGTACATACAGTGTCCCTCTCTTTGCTGAAA	492	
QY	580	AACCCCTGATTATTTTTCAGGAAGAGCAATACATATTTGTGTATTTGGGAAGCTTTC	639	
Db	493	AATCCTGACTATTTTTCAAAAGAAGCAAGCAGGAGCCATTTATGTATCTGGGGCCCTTTC	552	
QY	640	CGCAATATGAGAAAAGATGGCAATGGCATCGTTTACAACATGTTGAAAAGACAGTTCGT	699	
Db	553	CG-AACATGAGGAAGATGGTAATGGCATCGTGTACACATGTTTAAGAGACAGTTCAT	611	
QY	700	ATCTATCCGAATGCCAAATATAC	723	
Db	612	GCCTATCCAGATGCGCAGATCTAC	635	

RESULT 11
HSM076229
ID HSM076229 standard; RNA; EST; 330 BP.
XX
XX BX486042;
XX

SV	EX486042.1	SV
XX		
DT	09-MAY-2003 (Rel. 75, Created)	
DT	09-MAY-2003 (Rel. 75, Last updated, Version 1)	
DE	Homo sapiens mRNA; EST DKFZp686003249_r1 (from clone DKFZp686003249)	
XX		
XX	EST; expressed sequence tag.	
XX		
XX	Homo sapiens (human)	
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia	
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.	
XX		
XX	[1]	
RP	1-330	
RA	Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,	
RA	Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;	
RT	; Submitted (07-MAY-2003) to the EMBL/GenBank/DBSJ databases.	
RL	MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY	
XX		
XX	This is the 5' sequence of the clone insert	
CC	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
CC	Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;	
CC	sequenced by EMBL (European Molecular Biology Laboratories,	
CC	Heidelberg/Germany), within the cDNA sequencing consortium of	
CC	the German genome project.	
CC	No sl sequence available.	
CC	This clone (DKFZp686003249) is available at the RZPD in Berlin.	
CC	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,	
CC	14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de	
XX		
XX	Key	Location/Qualifiers
XX		
FT	source	1. 330
FT		/db_xref="taxon:9606"
FT		/mol_type="mRNA"
FT		/organism="Homo sapiens"
FT		/clone="DKFZp686003249"
FT		/clone_lib="686 (synonym: hlccc3). Vector pSport1_Sfi; host
FT		DH10B; sites SfiI + SfiII"
FT		/dev_stage="adult"
FT		/tissue_type="cDNA-collection"
XX		
XX	Sequence 330 BP; 71 A; 96 C; 91 G; 72 T; 0 other;	
XX		
QY	Query Match	24.3%; Score 314.2; DB 2; Length 330;
QY	Best Local Similarity	98.8%; Pred. No. 6.3e-82;
QY	Matches 327; Conservative	0; Mismatches 3; Indels 1; Gaps 1;
QY	52	CGGAAATGTGGGCTGGAGAGTCTCTGCGTGGTACAGCCCTCCAGCGTGCCTCCAGGACT 111
Db	1	CGGAAATGTGGGCTGGAGAGTCTCTGCGTGGTACAGCCCTCCAGCGTGCCTCCAGGACT 60
QY	112	GCCCTCGACCCAGGCGCGCCGCTGCTCGGTGGAGAGGCGCGGCGGAGCGCCATGCC 171
Db	61	GCCCTCGACCCAGGCGCGCCGCTGCTCGGTGGAGAGGCGCGGCGGAGCGCCATGCC 120
QY	172	TGCATCTCGAAGAGAAAGTCTGTGATTGTGTGAGCTTCATAGCAGGTTCTCTTTCCTG 231
Db	121	TGCATCTCGAAGAGAAAGTCTGTGATTGTGTGAGCTTCATAGCAGGTTCTCTTTCCTG 180
QY	232	CTGTTGTGCGTCTGTGTAAGTGAAGTGAATTTCCATTGCTTACTTAACATGCTTTGGCAA 291
Db	181	CTGTTGTGCGTCTGTGTAAGTGAAGTGAATTTCCATTGCTTACTTAACATGCTTTGGCAA 239
QY	292	CCTGGTACAAAGTGAATACCATTTCTCTACATACATAGCGCGCCCTTGGAACTCACTAT 351
Db	240	CCTGGTACAAAGTGAATACCATTTCTCTACATACATAGCGCGCCCTTGGAACTCACTAT 299
QY	352	GGATACATAAATGTGAAGACACAAAGCGCTT 382
Db	300	GGATACATAAATGTGAAGACACAAAGCGCTT 330

[illegible]

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 481)
AUTHORS	Okazaki,Y., Furuno,M., Sasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikiido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriani,L.M., Kaspin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Brad,T.D., Bruscia,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Giasi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sanderlin,A., Schneider,C., Sempke,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wyshaw-Boris,A., Yanagisawa,M., Yang,I., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,I., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Yeterstein,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	22354683
PUBMED	12466851
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res.gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane-Kishikawa,T., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Wataniki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES		Location/Qualifiers	
source		I. .481	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="C57BL/6J"	
		/db_xref="taxon:10090"	
		/clone="G370026007"	
		/cell_type="B16 F10Y cells"	
BASE COUNT		110 a	126 c 139 g 106 t
ORIGIN			
Query Match		17.7%	Score 228.4; DB 13; Length 481;
Best Local Similarity		79.4%	Pred. No. 2.6e-56;
Matches 308;		Conservative 0;	Mismatches 76; Indels 4; Gaps 3;
QY	160	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	219
DB	97	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	156
QY	220	TTCCTTTCCGCTGGTGTGCGCTCTGTAAATGAAGTAATTTCCCATGCTGCTACTAAAC	279
DB	157	TGCATCTTCTGCTAGCCATCGCGCTTGTAATGATGCGACTTTCCCTTTGCTCCTGAAC	216
QY	280	TGCTTTGGACAACCTGGTACAAAGTGTGATACCATCTCTACATACATAGCGGCGCC-CT	338
DB	217	TGCTTTGGAGCGCTAGACCAATAGATCCCTTTGCCCTACATTCAGGAGCGCCCTCT	276
QY	339	TGGAATCACTATGATACATAAATATGTAAGACACAGAGCGCTTTGCACTGGACTGTA	398
DB	277	TGGAATCACTATGATACATAAATATGTAAGACACAGAG-CTTTGGAAGTGAAGTGA	335
QY	399	CTTTTGGCACTAGTGTCAACATCAGGTGAGTGTGTCGCGCAGAGTGGGAAATGAGAT	458
DB	336	CCAGTGTGCACTAGTGTCAACATCAGGTGAGTGTGTCGCGCAGAGTGGGAAATGAGAT	395
QY	459	AGATCATCTCTCTGCATTTGGAGATGACAAATGCGCCCAACAAAGTTATGAAGAAGA	518
DB	396	AGACCATGATCTCTCATCTGGAGATGAGACA--CGCCCAACCAAGGCTTCGAGGACA	453
QY	519	TGTCGCCCGCAGTACCAATGATCGAGTT 546	
DB	454	TGTCGCCCGCAGTACCAATGATCGAGTT 481	

RESULT 15	
LOCUS	BY017303
DEFINITION	BY017303 RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA Mus musculus cDNA clone G830021L03 5', mRNA
ACCESSION	BY017303
VERSION	BY017303.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 386)
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nakaide,I., Saito,N., Saito,K., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldairelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki

FEATURES

source		Location/Qualifiers	
		I. .386	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/clone="G830021L03"	
		/tissue_type="mammary gland"	
		/cell_line="RCB-0526 Jyg-MC(A)"	
		/clone_lib="RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA"	
BASE COUNT		87 a	104 c 106 g 89 t
ORIGIN			
Query Match		17.1%	Score 221.6; DB 13; Length 386;
Best Local Similarity		81.3%	Pred. No. 2.5e-54;
Matches 257;		Conservative 0;	Mismatches 59; Indels 0; Gaps 0;
QY	160	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	219
DB	71	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	130

H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,C.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sempile,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Aizawa,K., Akimura,Y., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct

Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

I. .386		Location/Qualifiers	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/clone="G830021L03"	
		/tissue_type="mammary gland"	
		/cell_line="RCB-0526 Jyg-MC(A)"	
		/clone_lib="RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA"	
BASE COUNT		87 a	104 c 106 g 89 t
ORIGIN			
Query Match		17.1%	Score 221.6; DB 13; Length 386;
Best Local Similarity		81.3%	Pred. No. 2.5e-54;
Matches 257;		Conservative 0;	Mismatches 59; Indels 0; Gaps 0;
QY	160	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	219
DB	71	AGCGCCATGGCTGCATCTCAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCG	130

```
QY      220 TTCCTTTCCCTGCTGGTGTGGCGCTCTGTGAAATGAAGTGAANTTCCCAATGCTACTATAAC 279
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      131 TGCATCTTGTCTGTAGCCATGGCCCTTGTCAATGATCGACTTTCCTTTTGTCTCTGAAC 190
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      280 TGCTTTGGACAACCTGCTACAAAGTGGATACCATTCTCTACACATACAGGGGGCCCTT 339
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      191 TGCTTTGGACAGCCTAAGACCAATGGATGCCCTTTGCCCTACACATCAGGAGCCCTT 250
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      340 CGAACTCACTATGGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGGACTGTGAC 399
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      251 CGAACTCACTATGGATACATAAATGTGAAGACACAAGAGCCTTTGCACTGGACTGTGAC 310
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      400 CTTTGTGCCATAGTCTCAACTCAGGTCAGATGGTTGGCCAGAAAGTGGGAAATGAGATA 459
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      311 CACTGTGCCATCGTGTCAATTCAGGTCAAAATGTCGGGCAAGAGTGGGGGGAAGAGATA 370
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      460 GATCGATCCTCCTGCA 475
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      371 GACCATGCATCCTGCA 386
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: September 12, 2003, 15:02:54
Job time : 3114 secs